

# PATENT ABSTRACTS OF JAPAN

(11)Publication number : 06-250795

(43)Date of publication of application : 09.09.1994

(51)Int.CI.

G06F 3/06  
G06F 3/06

(21)Application number : 05-035233

(71)Applicant : HITACHI LTD

(22)Date of filing : 24.02.1993

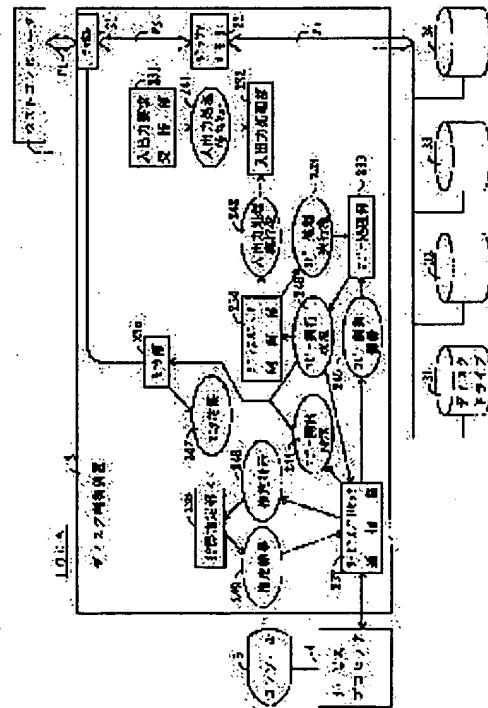
(72)Inventor : AKATSU MASAHIRO  
MURATA TOMOHIRO  
KURIHARA KENZO  
KURANO AKIRA

## (54) DISK ARRAY SYSTEM

### (57)Abstract:

PURPOSE: To provide the disk array system which shows an index for determining to what extent a copying processing is performed preferentially to an operator and performs optimum control according to the index when a disk restoring processing is performed.

CONSTITUTION: A monitor part 235 measures an input/output request arrival rate, a mean response time, and a disk restoration time and records them as a monitor result 236. At the time of start of disk restoration or during the disk restoration, an index prediction part 236 predicts a disk restoration time, a mean response time, and a mean performance maintaining time by using a current input/output arrival rate and the monitor result 236, and shows it to the operator as reference at the time of designating of a copying speed and also automatically performs optimum control on the basis of it to carry out the copying process. Consequently, the disk restoring processing from a viewpoint of both reliability and on-line processing performance can be performed.



## LEGAL STATUS

[Date of request for examination]

[Date of sending the examiner's decision of rejection]

[Kind of final disposal of application other than the examiner's decision of rejection or application converted registration]

[Date of final disposal for application]

[Patent number]

[Date of registration]

[Number of appeal against examiner's decision of rejection]

[Date of requesting appeal against examiner's decision of rejection]

[Date of extinction of right]

Copyright (C); 1998,2003 Japan Patent Office

JP06-250795

## \* NOTICES \*

Japan Patent Office is not responsible for any  
damages caused by the use of this translation.

1. This document has been translated by computer. So the translation may not reflect the original precisely.
2. \*\*\*\* shows the word which can not be translated.
3. In the drawings, any words are not translated.

---

CLAIMS

## [Claim(s)]

[Claim 1] When it consists of a host computer and a disk controller which are characterized by providing the following, and two or more disk drives, data is distributed and recorded on two or more disk drives and one or more sets of disk drives break down, Data holding a broken disk is restored based on data in a normal disk drive. A disk array system which performs restoration processing to said failure while continuing on-line processing by performing copy processing written in a disk drive fixed and exchanged or a disk drive of a spare at copy speed specified as an operator A monitor means to measure an input/output request arrival rate from a host computer, the disk release time to the completion of copy processing, copy speed, and an average input/output request arrival rate under disk restoration processing, and to record them as a monitor result An index presumption means to ask for copy speed when assuming that an input/output request arrival rate in this time was maintained, and relation of the disk release time based on said monitor result, and to output it to it as a presumed result A service means to display said presumed result as an index for copy speed assignment, and to make an operator specify copy speed

[Claim 2] When it consists of a host computer and a disk controller which are characterized by providing the following, and two or more disk drives, data is distributed and recorded on two or more disk drives and one or more sets of disk drives break down, Data holding a broken disk is restored based on data in a normal disk drive. A disk array system which performs restoration processing to said failure while continuing on-line processing by performing copy processing written in a disk drive fixed and exchanged or a disk drive of a spare at copy speed specified as an operator A monitor means to measure an input/output request arrival rate from a host computer, the disk release time to the completion of copy processing, copy speed, and an average input/output request arrival rate under disk restoration processing, and to record them as a monitor result An index presumption means to find the disk release time when assuming that an input/output request arrival rate in this time was maintained based on copy speed specified as an operator, and said monitor result, and to output it to it as a presumed result A service means to display said presumed result

[Claim 3] When it consists of a host computer, a disk controller, and two or more disk drives, data is distributed and recorded on two or more disk drives and one or more sets of disk drives break down, Data holding a broken disk is restored based on data in a normal disk drive. By performing copy processing written in a disk drive fixed and exchanged or a disk drive of a spare at copy speed specified as an operator In a disk array system which performs restoration processing to said failure while continuing on-line processing An input/output request arrival rate from a host computer, and the disk release time to the completion of copy processing, A monitor means to measure copy speed and an average input/output request arrival rate under disk restoration processing, and to record them as a monitor result, When it assumes that an input/output request arrival rate in this time was maintained based on said monitor result It asks for \*\* copy speed and relation of the disk release time. Further based on the disk release time, a failure rate of one set of a disk drive, and a rate of repair exchange by failure of two or more sets of disk units An index presumption means to calculate a mean time MTTDL

until data restoring becomes impossible, and to output it as a presumed result, A disk array system characterized by providing a service means to display said presumed result as an index for copy speed assignment, and to make an operator specify copy speed.

[Claim 4] When it consists of a host computer and a disk controller which are characterized by providing the following, and two or more disk drives, data is distributed and recorded on two or more disk drives and one or more sets of disk drives break down, Data holding a broken disk is restored based on data in a normal disk drive. A disk array system which performs restoration processing to said failure while continuing on-line processing by performing copy processing written in a disk drive fixed and exchanged or a disk drive of a spare at copy speed specified as an operator A monitor means to measure an input/output request arrival rate from a host computer, an average response time to the completion of processing of input/output request from a host computer of a under [ disk restoration processing ], copy speed, and an average input/output request arrival rate under disk restoration processing, and to record them as a monitor result An index presumption means to ask for copy speed when assuming that an input/output request arrival rate in this time was maintained, and relation of an average response time based on said monitor result, and to output it to it as a presumed result A service means to display said presumed result as an index for copy speed assignment, and to make an operator specify copy speed

[Claim 5] When it consists of a host computer, a disk controller, and two or more disk drives, data is distributed and recorded on two or more disk drives and one or more sets of disk drives break down, Data holding a broken disk is restored based on data in a normal disk drive. By performing copy processing written in a disk drive fixed and exchanged or a disk drive of a spare at copy speed specified as an operator In a disk array system which performs restoration processing to said failure while continuing on-line processing An input/output request arrival rate from a host computer, and the response time to the completion of processing of input/output request from a host computer, A monitor means to measure the disk release time to the completion of copy processing, copy speed, and an average input/output request arrival rate under disk restoration processing, and to record them as a monitor result, When it assumes that an input/output request arrival rate in this time was maintained based on said monitor result When it assumes that an input/output request arrival rate in this time was maintained based on the aim response time which asked for \*\* copy speed and relation of the disk release time, and was specified as an operator, and said monitor result The response time to \*\*\*\* copy speed A condition classification which presumes, makes an engine-performance maintenance condition when the response time is smaller than the aim response time, and is made into a performance degradation condition when large is performed. Based on said disk release time, a failure rate of one set of a disk drive, a rate of repair exchange, and said condition classification by failure of two or more sets of disk drives Average engine-performance maintenance time amount which will function in the state of said engine-performance maintenance by the time data restoring becomes impossible is calculated. A disk array system characterized by providing an index presumption means to output it as a presumed result, and a service means to display said presumed result as an index for copy speed assignment, and to make an operator specify copy speed.

[Claim 6] In a disk array system according to claim 5 from claim 1 a monitor means A copy processing activation condition is checked and it records on a monitor result. An index presumption means During copy processing activation, residual time which disk restoration takes based on said monitor result is presumed, and it records on a presumed result. A service means During copy processing activation, display said presumed result as an index for copy speed modification directions, and an operator is made to do the modification directions of the copy speed. Furthermore, a disk array system characterized by providing a copy processing means to change copy speed and to perform copy processing when there are copy speed modification directions from a copy processing operator.

[Claim 7] From claim 1 characterized by providing the following to a disk array system according to claim 5 A monitor means checks a copy processing activation condition, and records it on a monitor result. A service means From an operator the aim response time to reception and a pan During copy processing activation, said monitor result and said aim response time When an average response time is larger than the aim response time, when an average response time is smaller than the aim response time,

change copy speed more greatly, or change frequency of copy activation more highly, change copy speed into a basis more smallish, or or frequency of copy activation A copy adjustment means to change lowness A copy processing means to change copy conditions and to perform copy processing when there is modification of copy conditions by the copy adjustment means

[Claim 8] A disk array system characterized by starting actuation for copy processing of each means, without waiting for directions of an operator when it has a disk drive of a spare and an aim speed of response and copy speed at the time of copy initiation are beforehand given in a disk array system according to claim 7.

[Claim 9] When it consists of a host computer, a disk controller, and two or more disk drives, data is distributed and recorded on two or more disk drives and one or more sets of disk drives break down, Data holding a broken disk is restored based on data in a normal disk drive. Copy processing written in a disk drive fixed and exchanged or a disk drive of a spare by performing at a predetermined copy speed In a disk array system which performs restoration processing to said failure while continuing on-line processing An input/output request arrival rate from a host computer, and the response time to the completion of processing of input/output request from a host computer, A monitor means to measure the disk release time to the completion of copy processing, copy speed, and an average input/output request arrival rate under disk restoration processing, and to record them as a monitor result, When it assumes that an input/output request arrival rate in this time was maintained based on said monitor result When it assumes that an input/output request arrival rate in this time was maintained based on the aim response time which asked for \*\* copy speed and relation of the disk release time, and was specified as an operator, and said monitor result The response time to \*\*\*\* copy speed It presumes. A condition classification which makes an engine-performance maintenance condition when the average response time is smaller than the aim response time, and is made into a performance degradation condition when large is performed. Based on said disk release time, a failure rate of one set of a disk drive, a rate of repair exchange, and said condition classification by failure of two or more sets of disk drives Average engine-performance maintenance time amount which will function in the state of said engine-performance maintenance by the time data restoring becomes impossible is calculated. An index presumption means to output it as a presumed result, and a copy adjustment means to set up copy speed from which average engine-performance maintenance time amount serves as max with reference to said presumed result, A disk array system characterized by providing a copy processing means to start copy processing with copy speed set up by the copy adjustment means.

[Claim 10] In a disk array system according to claim 9 a monitor means While measuring the response time, elapsed time from the number of trucks and copy initiation which ended a copy is measured, and it is set as a monitor result. An index presumption means Average engine-performance maintenance time amount over each copy speed is re-presumed based on the aim response time specified by operator during copy processing activation, and a monitor result, and it is set as a presumed result. A copy adjustment means It is the disk array system which changes into copy speed from which average engine-performance maintenance time amount serves as max with reference to said presumed result a suitable period, and is characterized by changing into the copy speed and performing copy processing when a copy means has modification of copy speed with said copy adjustment means.

[Claim 11] A disk array system characterized by starting actuation for copy processing of each means, without waiting for directions of an operator when it has a disk drive of a spare and an aim speed of response is beforehand given in a disk array system according to claim 9 or 10.

---

[Translation done.]

## \* NOTICES \*

Japan Patent Office is not responsible for any damages caused by the use of this translation.

1. This document has been translated by computer. So the translation may not reflect the original precisely.
2. \*\*\*\* shows the word which can not be translated.
3. In the drawings, any words are not translated.

---

## DETAILED DESCRIPTION

---

### [Detailed Description of the Invention]

[0001]

[Industrial Application] This invention relates to the disk array system which can perform suitably disk restoration processing when a disk drive breaks down in more detail about a disk array system.

[0002]

[Description of the Prior Art] When it distributes to two or more disk drives, a disk array system records data and one or more sets of disk drives break down, The data holding the broken disk drive is restored based on the data in a normal disk drive. About the copy processing written in the disk drive fixed and exchanged or the disk drive of a spare, it is the system which enabled it to perform restoration processing to said failure a number truck every, continuing on-line processing by carrying out. Such a disk array system is indicated by commercial production \*\*\*\* disk array (1992.3.9; 87-98 pages of Nikkei computers)" towards "spread.

[0003]

[Problem(s) to be Solved by the Invention] In a disk array system, if the priority of copy processing is raised, since the time amount which has stopped at the condition that some disk drives are out of order will become short, the probability which lapses into the condition that a data reload becomes impossible according to a multiplex failure falls. That is, reliability improves. On the other hand, since on-line processing becomes whether to be a non-dense only in the part, the on-line-processing engine performance falls. That is, although on-line processing and copy processing are performed to juxtaposition in a disk array system, it is difficult to determine whether to give priority to copy processing how much compared with on-line processing, and to perform it. However, the technology about the priority of copy processing is not indicated conventionally.

[0004] It is in offering the disk array system which the index for determining to copy processings [ how many ] priority should be given in case disk restoration processing is performed is calculated , an operator is shown , or the place made into the purpose performs [ this invention was made in view of such a situation , ] optimal control based on it , and performs suitable disk restoration processing from a viewpoint of both reliability and the on-line processing engine performance .

[0005]

[Means for Solving the Problem] In the 1st viewpoint, this invention An input/output request arrival rate from a host computer, A monitor means to measure the disk release time to the completion of copy processing, copy speed, and an average input/output request arrival rate under disk restoration processing, and to record them as a monitor result, An index presumption means to ask for copy speed when assuming that an input/output request arrival rate in this time was maintained, and relation of the disk release time based on said monitor result, and to output it to it as a presumed result, Said presumed result is displayed as an index for copy speed assignment, and a disk array system characterized by providing a service means to make an operator specify copy speed is offered.

[0006] In the 2nd viewpoint, this invention An input/output request arrival rate from a host computer, A monitor means to measure the disk release time to the completion of copy processing, copy speed, and

an average input/output request arrival rate under disk restoration processing, and to record them as a monitor result, An index presumption means to find the disk release time when assuming that an input/output request arrival rate in this time was maintained based on copy speed specified as an operator, and said monitor result, and to output it to it as a presumed result, A disk array system characterized by providing a service means to display said presumed result is offered.

[0007] In the 3rd viewpoint, this invention An input/output request arrival rate from a host computer, A monitor means to measure the disk release time to the completion of copy processing, copy speed, and an average input/output request arrival rate under disk restoration processing, and to record them as a monitor result, When it assumes that an input/output request arrival rate in this time was maintained based on said monitor result It asks for \*\* copy speed and relation of the disk release time. Further based on the disk release time, a failure rate of one set of a disk drive, and a rate of repair exchange by failure of two or more sets of disk units An index presumption means to calculate the mean time MTTDL until data restoring becomes impossible, and to output it as a presumed result, Said presumed result is displayed as an index for copy speed assignment, and a disk array system characterized by providing a service means to make an operator specify copy speed is offered.

[0008] In the 4th viewpoint, this invention An input/output request arrival rate from a host computer, An average response time to the completion of processing of input/output request from a host computer of a under [ disk restoration processing ], A monitor means to measure copy speed and an average input/output request arrival rate under disk restoration processing, and to record them as a monitor result, An index presumption means to ask for copy speed when assuming that an input/output request arrival rate in this time was maintained, and relation of an average response time based on said monitor result, and to output it to it as a presumed result, Said presumed result is displayed as an index for copy speed assignment, and a disk array system characterized by providing a service means to make an operator specify copy speed is offered.

[0009] In the 5th viewpoint, this invention An input/output request arrival rate from a host computer, The response time to the completion of processing of input/output request from a host computer, A monitor means to measure the disk release time to the completion of copy processing, copy speed, and an average input/output request arrival rate under disk restoration processing, and to record them as a monitor result, When it assumes that an input/output request arrival rate in this time was maintained based on said monitor result When it assumes that an input/output request arrival rate in this time was maintained based on the aim response time which asked for \*\* copy speed and relation of the disk release time, and was specified as an operator, and said monitor result The response time to \*\*\*\* copy speed A condition classification which presumes, makes an engine-performance maintenance condition when the response time is smaller than the aim response time, and is made into a performance degradation condition when large is performed. Based on said disk release time, a failure rate of one set of a disk drive, a rate of repair exchange, and said condition classification by failure of two or more sets of disk drives Average engine-performance maintenance time amount which will function in the state of said engine-performance maintenance by the time data restoring becomes impossible is calculated. A disk array system characterized by providing an index presumption means to output it as a presumed result, and a service means to display said presumed result as an index for copy speed assignment, and to make an operator specify copy speed is offered.

[0010] In the 6th viewpoint, this invention is set to a disk array system of the above-mentioned configuration. A monitor means A copy processing activation condition is checked and it records on a monitor result. An index presumption means During copy processing activation, residual time which disk restoration takes based on said monitor result is presumed, and it records on a presumed result. A service means During copy processing activation, display said presumed result as an index for copy speed modification directions, and an operator is made to do the modification directions of the copy speed. Furthermore, when there are copy speed modification directions from a copy processing operator, a disk array system characterized by providing a copy processing means to change copy speed and to perform copy processing is offered.

[0011] In the 7th viewpoint, this invention is set to a disk array system of the above-mentioned

configuration. A monitor means A copy processing activation condition is checked and it records on a monitor result. A service means From an operator the aim response time to reception and a pan During copy processing activation, said monitor result and said aim response time When an average response time is larger than the aim response time, when an average response time is smaller than the aim response time, change copy speed more greatly, or change frequency of copy activation more highly, change copy speed into a basis more smallish, or or frequency of copy activation A disk array system characterized by providing a copy adjustment means to change lowness, and a copy processing means to change copy conditions and to perform copy processing when there is modification of copy conditions by the copy adjustment means is offered.

[0012] In the 8th viewpoint, in a disk array system by the 7th viewpoint of the above, this invention offers a disk array system characterized by starting actuation for copy processing of each means, without waiting for directions of an operator, when it has a disk drive of a spare and the aim response time and copy speed at the time of copy initiation are given beforehand.

[0013] In the 9th viewpoint, this invention An input/output request arrival rate from a host computer, The response time to the completion of processing of input/output request from a host computer, A monitor means to measure the disk release time to the completion of copy processing, copy speed, and an average input/output request arrival rate under disk restoration processing, and to record them as a monitor result, When it assumes that an input/output request arrival rate in this time was maintained based on said monitor result When it assumes that an input/output request arrival rate in this time was maintained based on the aim response time which asked for \*\* copy speed and relation of the disk release time, and was specified as an operator, and said monitor result The response time to \*\*\*\* copy speed It presumes. A condition classification which makes an engine-performance maintenance condition when the average response time is smaller than the aim response time, and is made into a performance degradation condition when large is performed. Based on said disk release time, a failure rate of one set of a disk drive, a rate of repair exchange, and said condition classification by failure of two or more sets of disk drives Average engine-performance maintenance time amount which will function in the state of said engine-performance maintenance by the time data restoring becomes impossible is calculated. An index presumption means to output it as a presumed result, and a copy adjustment means to set up copy speed from which average engine-performance maintenance time amount serves as max with reference to said presumed result, A disk array system characterized by providing a copy processing means to start copy processing with copy speed set up by the copy adjustment means is offered.

[0014] In a disk array system according [ this invention ] to the 9th viewpoint of the above at the 10th viewpoint a monitor means While measuring the response time, elapsed time from the number of trucks and copy initiation which ended a copy is measured, and it is set as a monitor result. An index presumption means Average engine-performance maintenance time amount over each copy speed is re-presumed based on the aim response time specified by operator during copy processing activation, and a monitor result, and it is set as a presumed result. A copy adjustment means It changes into copy speed from which average engine-performance maintenance time amount serves as max with reference to said presumed result a suitable period, and a copy means provides a disk array system characterized by changing into the copy speed and performing copy processing, when there is modification of copy speed with said copy adjustment means.

[0015] In the 11th viewpoint, in a disk array system by the 9th viewpoint of the above, or the 10th viewpoint, this invention offers a disk array system characterized by starting actuation for copy processing of each means, without waiting for directions of an operator, when it has a disk drive of a spare and an aim speed of response is given beforehand.

[0016]

[Function] In the disk array system of this invention by the 1st viewpoint of the above, the copy speed when assuming that the input/output request arrival rate in this time was maintained as an index for determining to copy processings [ how many ] priority should be given and the relation of the disk release time are shown to an operator. Then, an operator can specify copy speed which serves as the

suitable disk release time from a viewpoint of both reliability and the on-line-processing engine performance.

[0017] In the disk array system of this invention by the 2nd viewpoint of the above, copy processing is carried out at the copy speed specified as the operator, and the disk release time when assuming that the input/output request arrival rate in this time was maintained is shown to an operator. Then, an operator can check whether copy speed which serves as the suitable disk release time from a viewpoint of both reliability and the on-line-processing engine performance has been specified.

[0018] In the disk array system of this invention by the 3rd viewpoint of the above, the mean time MTTDL until data restoring becomes impossible as an index for determining to copy processings [ how many ] priority should be given is shown to an operator. Then, an operator can specify copy speed which serves as the suitable disk release time from a viewpoint of both reliability and the on-line-processing engine performance.

[0019] In the disk array system of this invention by the 4th viewpoint of the above, the copy speed when assuming that the input/output request arrival rate in this time was maintained as an index for determining to copy processings [ how many ] priority should be given and the relation of an average response time are shown to an operator. Then, an operator can specify copy speed which serves as a suitable average response time from a viewpoint of both reliability and the on-line-processing engine performance.

[0020] When calling the total time amount which will function in the state of said engine-performance maintenance by the time it becomes impossible an engine-performance maintenance condition, a call, and to data restore about the case where the response time to each copy speed is smaller than the aim response time average engine-performance maintenance time amount, in the disk array system of this invention by the 5th viewpoint of the above, an operator shows as an index for determining whether to copy processings [ how many ] to give priority for the average engine-performance maintenance time amount. Then, an operator can specify copy speed which serves as suitable average engine-performance maintenance time amount from a viewpoint of both reliability and the on-line-processing engine performance.

[0021] By the disk array system of this invention by the 6th viewpoint of the above, the residual time which disk restoration takes is presumed during copy processing activation, it is shown to an operator, the modification directions of the copy speed are carried out, and copy processing is continued at the copy speed which the copy processing operator changed. Then, when an input/output request arrival rate is changed during copy processing activation, it can respond.

[0022] In the disk array system of this invention by the 7th viewpoint of the above, the automatic regulation of copy speed or the frequency of copy activation is carried out so that an average response time may be in agreement during copy processing activation at the aim response time. Then, when an input/output request arrival rate is changed during copy processing activation, it can respond.

[0023] By the disk array system of this invention by the 8th viewpoint of the above, when it has the disk drive of a spare, the actuation for copy processing of each means is started at the copy speed at the time of the copy initiation given beforehand, without waiting for directions of an operator. Then, directions of an operator become unnecessary and maintenance can be automated.

[0024] In the disk array system of this invention by the 9th viewpoint of the above, copy speed is set automatically and copy processing is started so that average engine-performance maintenance time amount may become max. Then, directions of an operator become unnecessary. In the disk array system of this invention by the 10th viewpoint of the above, during copy processing activation, the automatic regulation of the copy speed is carried out so that average engine-performance maintenance time amount may become max. Then, when an input/output request arrival rate is changed during copy processing activation, it can respond. By the disk array system of this invention by the 11th viewpoint of the above, when it has the disk drive of a spare, the actuation for copy processing of each means is started at the copy speed at the time of the copy initiation given beforehand, without waiting for directions of an operator. Then, directions of an operator become unnecessary and maintenance can be automated.

[0025]

[Example] Hereafter, a drawing explains the example of this invention to details. In addition, thereby, this invention is not limited. Drawing 1 is the block diagram of the hardware configuration common to the disk array system applied to the 9th example from the 1st example mentioned later. This disk array system 100 consists of a host computer 1, a disk controller 2, disk drives 31-34, a service processor 4, and a console 5. In this example, data shall be distributed and written in three of four sets of disk drives 31-34, and an error correction code shall be written in the one remaining sets. However, in this invention, there is no constraint in the number of a disk drive,  $m$  sets of disk drives are prepared, and it is good as a drive for error corrections in  $n$  of sets [ them ] (natural number of arbitration with which  $m$  and  $n$  fill  $m > n \geq 1$ ). The disk controller 2 consists of a channel 21, buffer memory 22, control processors 25 and 26, local memories 27 and 28, and a common memory 29. As for the 9th example, software configurations, i.e., a function, differ from the 1st example, respectively. Next, the 9th example is explained in order from the 1st example.

[0026] - 1st example- drawing 2 is the block diagram of disk array system 100A concerning the 1st example of this invention. A disk controller 2 has a channel 21 and buffer memory 22. The data transfer ways P1, P2, and P3 are between buffer memory 22 and disk drives 31-34 between a channel 21 and buffer memory 22 between a host computer 1 and a channel 21, respectively.

[0027] It is square and surrounded each part 231-237 in a disk controller 2 expresses the function realized by the micro program of control processors 25 and 26. In addition, one control processor may be assigned for every function, and all functions may be given only to one control processor. Moreover, each information 241-249 surrounded with the ellipse in a disk controller 2 is information which delivers between each above-mentioned function. The information which delivers between the functions currently assigned to different control processors 25 and 26 among such information is placed into the common memory 29 which each control processors 25 and 26 can refer to and update. Moreover, the information which delivers between the functions currently assigned to the same control processor is put on the local memories 27 or 28 of the control processor.

[0028] When a host computer 1 publishes input/output request (a data read-out instruction / data write-in instruction) to disk drives 31-34 through a channel 21, the input/output request acceptance section 231 detects it, and registers it into the waiting queue 241 for radial transfer.

[0029] The radial transfer section 232 processes the input/output request registered into the head of the waiting queue 241 for radial transfer, when the right 242 of radial transfer activation is granted from the dispatch control section 234. If three sets of the disk drives which store data are normal when input/output request is a data read-out instruction, three sets of the disk drives will be accessed, and data will be read. And if one set is out of order, after restoring the data which reads data and the code for error corrections from remaining three sets, and is stored in the broken disk drive, it transmits to a host computer 1 via buffer memory 22. When input/output request is a data write-in instruction, while dividing data into three, the code for error corrections is created and each is written in a normal disk drive via buffer memory 22.

[0030] The copy processing section 233 performs copy processing, after fixing and exchanging them, when one of four sets of disk drives breaks down. Flow drawing of copy processing is shown in drawing 3 . First, a loop is carried out until it receives the copy initiation directions 244 from the service-processor communications department 237 (step 1000). If the copy initiation directions 244 are received, the copy control information 245 which consists of the disk drive number for restoration and the number of trucks (it is hereafter called copy speed.) treated by one copy processing will be read (step 1010). Next, the track number which performs copy processing is initialized (step 1020), and a loop is carried out until the right 243 of copy processing activation is granted from the dispatch control section 234 (step 1030).

[0031] If the right 243 of copy processing activation is granted, the track number of the head which will perform copy processing from now on will be read (step 1040). Next, data is read from three sets of normal disk drives, and it stores in buffer memory 22 (step 1050). The trucks which read data are a number of trucks which began from the above-mentioned track number and were directed at the above-mentioned copy speed.

[0032] Next, from the data read to buffer memory 22, the data of the disk drive for restoration is restored and it stores in buffer memory 22 (step 1060). And the restored data is transmitted to the disk drive for restoration from buffer memory 22 (step 1070).

[0033] Next, the head track number of a next copy processing object is updated (step 1080). Finally it judges whether disk restoration processing was completed (step 1090), and if it has not completed, it returns to said step 1030. If it completes, the completion of copy processing will be set as the copy activation condition 246 (step 1100), and it will return to said step 1000. In addition, the radial transfer section 232 resets the right 242 of radial transfer activation, after detecting the right 242 of radial transfer activation. Moreover, the copy processing section 233 resets the right 243 of copy processing activation, after detecting the right 243 of copy processing activation.

[0034] The dispatch control section 234 is suitable schedule processing, and grants the right 242 of radial transfer activation to the radial transfer section 232. Moreover, the right of copy processing activation is granted to the copy processing section 233. Flow drawing of schedule processing is shown in drawing 4. This schedule processing usually sets up the right 242 of radial transfer activation, and sets up the right of copy processing activation for every fixed period. First, the right 242 of radial transfer activation is set up until it receives the copy initiation directions 244 from the service-processor communications department 237 (steps 2000 and 2010).

[0035] If the copy initiation directions 244 are received from the service-processor communications department 237, a timer will be reset zero times (step 2020). And if a timer is checked and it has not gone through the predetermined setup time (for example, 50ms), the right 242 of radial transfer activation is set up (steps 2030 and 2040). If the predetermined setup time passes, the right 243 of copy processing activation will be set up (step 2050).

[0036] And if the completion of copy processing is not set as the copy activation condition 246 and it is returned and set as said step 2020, it will return to said step 2000 (step 2060). In addition, in setting processing (steps 2010, 2040, and 2050) of each right 242,243 of processing activation, a loop is carried out on that spot until the right of processing activation set as last time is reset.

[0037] Close always checks the coming input/output request from a host computer 1 to a channel 21, and the monitor section 235 measures an input/output request arrival rate. And the input/output request arrival rate is recorded on the monitor result 247. Moreover, the disk release time after the copy initiation directions 244 are set up until the completion of copy processing is set as the copy activation condition 246, copy speed, and the average input/output request arrival rate under disk restoration processing are recorded on the monitor result 247. Whenever disk restoration processing is performed, the monitor of these disk release time, copy speed, and the average input/output request arrival rate is carried out, statistics processing is carried out and relation as shown in drawing 5 is obtained.

[0038] From the service-processor communications department 237, the index presumption section 236 reads the input/output request arrival rate in the time for the presumed directions 248 from the monitor result 247 at the time of a carrier beam, asks for the copy speed when assuming that the input/output request arrival rate was maintained, and the relation of the disk release time from the relation of drawing 5, and sets the result as the presumed result 249. The copy speed in a certain input/output request arrival rate and the relation of the disk release time come to be shown in drawing 6.

[0039] The service-processor communications department 237 has the following three functions. If the presumed directions 248 of the disk release time are set as it and the presumed result 249 is set as it when disk restoration processing initiation is directed in the first place from a service processor 4, it will be notified to a service processor 4. Carrier beams, and the copy initiation directions 244 and the copy control information 245 are set [ second ] up for assignment of copy control information (the disk drive for restoration, copy speed) from a service processor 4. If the completion of copy processing is set as the copy activation condition 246 by the third, the completion of disk restoration processing will be notified to a service processor 4.

[0040] A service processor 4 displays the information sent from the disk controller 2 on console equipment 5 while telling the directions which the operator inputted through console equipment 5 to a disk controller 2.

[0041] Drawing 7 is the sequence diagram of disk restoration processing in which each above-mentioned function was used. This is notified to a disk controller 2 that a service processor 4 receives disk restoration directions from an operator (step 3010). (step 3000)

[0042] On the other hand, with a disk controller 2, after presuming the disk release time over each copy speed in the index presumption section 236 (step 3020), it is reported to a service processor 4 (step 3030). In a service processor 4, the presumed result of the disk release time over each copy speed is displayed on console equipment 5 (step 3040). And an operator's assignment of a desired copy speed notifies it to a disk controller 2 (step 3050).

[0043] In a disk controller 2, the copy processing section 233 performs copy processing according to the notified copy speed (step 3060). Disk restoration processing termination is reported to a service processor 4 after the completion of copy processing (step 3070). A service processor 4 displays the message of the completion of copy processing on console equipment 5, and tells an operator about it (step 3080).

[0044] In addition, when copy speed is specified at the same time an operator directs disk restoration processing, the index presumption section 236 sets the estimate of the disk release time over the specified copy speed as the presumed result 249.

[0045] Since according to the above example [ 1st ] the estimate of the disk release time over each copy speed is displayed when an operator directs disk restoration processing, it is effective in it becoming easy to form a disk restoration work plan that an operator can specify copy speed in consideration of the probability for the 2nd set of disk drives to break down, for example etc. In addition, the function of the index presumption section 236 or the monitor section 235 may be moved to a service processor 4 or a host computer 1, or the measuring machine of the dedication which has the function of the monitor section 235 may be formed.

[0046] - Although the disk release time was presumed and displayed in the 2nd example-1st example, presume and display a mean time until a data reload becomes impossible by failure of two or more sets of disk drives (them are two sets in the case of this example) in the 2nd example. This is realized by changing the function of the index presumption section 236 of drawing 2 as follows. In the index presumption section 236, the state transition diagram of a disk array system is analyzed. The state transition diagram in the case of this example becomes like drawing 8 . That is, it has the following four conditions and six transition.

A condition S0 Four sets of disk drives during operation normally : A condition S1 One set of a disk drive is breaking down. During operation normally [ the remainder ] by three sets : A condition S2 : It is [ 1 set disk drive's restoration processing ] under activation, and is a condition S3 during operation normally [ at least the three remaining sets ]. : By failure of two sets of disk drives, a data reload is impossible (system down).

Transition One of four sets of disk drives : Failure transition (S0, S1) : (S1, S2) The broken disk drive Repair / exchange transition (S2, S0) : -- completion transition of disk drive restoration processing (S2, S1): -- the disk drive for restoration -- failure transition (S1, S3): -- one of three sets of normal disk drives -- failure transition (S2, S3): -- one of three sets of normal disk drives by failure drawing 8 The label on each transition expresses a transition rate. lambda expresses the failure rate of one set of a disk drive. mu expresses the rate of repair exchange of one set of a disk drive. T expresses the disk release time. The value of a failure rate lambda and the rate mu of repair exchange is beforehand given to the system. An assumption of that these follow exponential distribution forms the following differential equation by the theory of Markov process. In addition,  $P_i(t)$  is probability which exists in Condition  $S_i$  in time of day  $t$ .

[0047]

[Equation 1]

$$\left. \begin{aligned}
 \frac{dP_0(t)}{dt} &= -4\lambda P_0(t) + \frac{1}{T} P_2(t) \\
 \frac{dP_1(t)}{dt} &= 4\lambda P_0(t) - (3\lambda + \mu) P_1(t) + \lambda P_2(t) \\
 \frac{dP_2(t)}{dt} &= \mu P_1(t) - (4\lambda + \frac{1}{T}) P_2(t) \\
 \frac{dP_3(t)}{dt} &= 3\lambda P_1(t) + 3\lambda P_2(t)
 \end{aligned} \right\} \cdots \text{ (数1)}$$

[0048] In the index presumption section 236, from the service-processor communications department 237, the input/output request arrival rate in the time is read for the presumed directions 248 from the monitor result 247 at the time of a carrier beam, and the disk release time over each copy speed when assuming that the input/output request arrival rate was maintained is found from the relation of drawing 5. Next, the value of lambda, mu, and T is assigned above (several 1), and a differential equation is solved. And mean time MTTDL until a data reload becomes impossible (Mean Time To Data Lost) It asks by the degree type and each copy speed and the relation of a mean time MTTDL which were obtained are set as the presumed result 249.

[0049]

[Equation 2]

$$MTTDL = \int_0^{\infty} P_3(t) dt \cdots \text{ (数2)}$$

[0050] Each copy speed and the relation of a mean time MTTDL are displayed on console equipment 5 like the 1st example.

[0051] Since according to the above example [ 2nd ] the estimate of a mean-time MTTDL value until the data reload to each copy speed becomes impossible is displayed when an operator directs disk restoration processing, it is effective in it becoming easy to form a disk restoration work plan that an operator can specify copy speed in consideration of it etc.

[0052] - Although the estimate of the disk release time was displayed and the estimate of a mean-time MTTDL value until a data reload becomes impossible was displayed in the 2nd example in the 3rd example-1st example, presume and display the response time after receiving a demand until it finishes processing in the 3rd example. The configuration of the 3rd example is the same as the 1st example and the 2nd example except the function of the monitor section 235 and the index presumption section 236.

[0053] In the monitor section 235, the monitor of the signal which passes along a channel 21 is always carried out, and the input/output request from a host computer 1 and the processing termination report which the radial transfer section 232 performs to a host computer 1 are checked. And an input/output request arrival rate and the response time after receiving a demand until it finishes processing are measured. The input/output request arrival rate and the response time which were measured are recorded on the monitor result 247. Furthermore, in the monitor section 235, an average input/output request arrival rate and an average response time are recorded at every one disk restoration processing together with the copy speed at that time at the monitor result 247. Statistics processing of copy speed, an average input/output request arrival rate, and the average response time is carried out, and relation as shown in drawing 9 is obtained.

[0054] From the service-processor communications department 237, the index presumption section 236 reads the input/output request arrival rate in the time for the presumed directions 248 of an average response time from the monitor result 247 at the time of a carrier beam, and asks for the relation between the copy speed when assuming that the input/output request arrival rate was maintained, and an average response time. And the result is set as the presumed result 249. The copy speed in a certain input/output request arrival rate and the relation of an average response time come to be shown in

drawing 10.

[0055] Since according to the above example [ 3rd ] the estimate of the response time to each copy speed is displayed when an operator directs disk restoration processing, in consideration of the on-line-processing engine performance under disk restoration processing, an operator can specify copy speed, and it is effective in becoming easy to form a disk restoration work plan.

[0056] - The disk release time (reliability) and an average response time (on-line-processing engine performance) have the relation of a trade-off to copy speed so that 4th example- drawing 6 and drawing 10 may be compared and understood. Although the operator specified copy speed abstractly in consideration of these balance in the 1st example - the 3rd example, the index called the average engine-performance maintenance time amount which expresses quantitatively the balance of reliability and the on-line-processing engine performance is presumed and displayed in the 4th example.

[0057] In the system in which the degeneracy and restoration like a disk array system are possible, the condition (performance degradation condition) that it is less than the aim engine performance with the condition (engine-performance maintenance condition) of exceeding and working is in the inside in the condition of working. Then, the total time amount as which the system was functioning in the state of engine-performance maintenance by the time it repeated the engine-performance maintenance condition and the performance degradation condition by turns and it became impossible to return to an engine-performance maintenance condition finally is considered as one index. This index is called the average engine-performance maintenance time amount THP (Total High Performance Time). The average engine-performance maintenance time amount THP is computed by the degree type. In addition, SH expresses the set of an engine-performance maintenance condition.

[0058]

[Equation 3]

$$THP = \int_0^{\infty} \sum_{s_i \in SH} P_i(t) dt \quad \cdots \text{ (数3)}$$

[0059] In the monitor section 235, the monitor of an input/output request arrival rate, and the response time and the disk release time in said each system states S0, S1, S2, and S3 is carried out, and it records on the monitor result 247.

[0060] In the index presumption section 236, the aim response time is received from the service-processor communications department 237 for the presumed directions 248 of the average engine-performance maintenance time amount THP to coincidence at the time of a carrier beam. This aim response time is collectively inputted, when an operator directs disk restoration to a service processor 4. Next, in the index presumption section 236, the response time to each copy speed in each system states S0, S1, S2, and S3 when assuming that the input/output request arrival rate in the time was maintained based on the monitor result 247 is presumed. And each system states S0, S1, S2, and S3 are classified into an engine-performance maintenance condition or a performance degradation condition by comparing the response time to each of that presumed copy speed with said aim response time. Next, in the index presumption section 236, the disk release time over the input/output request arrival rate in the time is found based on the monitor result 247. Next, the value of lambda, mu, and T is assigned above (several 1), and a differential equation is solved. Furthermore, the average engine-performance maintenance time amount THP over each copy speed is calculated from the above (several 3) using the result and said classification result.

[0061] Finally, in the index presumption section 236, copy speed and the relation of the average engine-performance maintenance time amount THP are set as the presumed result 249.

[0062] Since according to the above example [ 4th ] the estimate of the average engine-performance maintenance time amount THP over each copy speed is displayed when an operator directs disk restoration processing, from now on, the balance of reliability and the on-line-processing engine performance will be grasped quantitatively, an operator can specify copy speed, and it is effective in becoming easy to form a disk restoration work plan.

[0063] - The 5th example-5th example is an example in which the operator enabled it to change copy

speed during disk restoration processing.

[0064] The number of trucks which finished the copy whenever one copy processing finished it with the copy activation condition 246 as the copy processing section 233 in addition to setting up the completion of copy processing, when copy processing was completed is set as the copy activation condition 246. In the monitor section 235, the copy activation condition 246 is always checked and the number of trucks which ended the copy is set as the monitor result 247 together with the elapsed time from copy initiation.

[0065] In the index presumption section 236, residual time until it judges the activation condition from copy initiation and copy processing is completed from the monitor result 247 is presumed. And the presumed residual time is set as the presumed result 249. This processing may always be performed spontaneously, and when it detects that there was a demand from an operator by referring to the presumed directions 248, it may be performed.

[0066] In the service-processor communications department 237, if copy speed is anew directed by the operator while notifying the presumed result 249 to a service processor 4, it will be set as the copy control information 245. In the copy processing section 233, whenever it acquires the right 243 of copy processing activation and performs copy processing, copy speed is read from the copy control information 245, and copy processing is performed at the directed copy speed.

[0067] Since according to the above example [ 5th ] copy speed can be changed while residual time until copy processing is completed is displayed during disk restoration processing, even when an input/output request arrival rate changes during disk restoration processing, it is effective in the ability of an operator to cope with it flexibly.

[0068] - The 6th example-6th example is an example which carries out the monitor of the response time and carries out the automatic regulation of the copy speed. Drawing 11 is the block diagram showing disk array system 100B of the 6th example. In the monitor section 235, the response time is always measured and it is set as the monitor result 247. In the service-processor communications department 237, disk restoration directions are combined from a service processor 4 at the time of a carrier beam, and the aim response time is received. And while setting the aim response time as desired value 250, the copy initiation directions 244 are set up.

[0069] By the copy controller 238, it is always being begun from the monitor result 247 to read the response time, and it is compared with desired value 250. Since it is possible to give priority to copy processing more when the desired value is larger, a bigger copy speed is set as the copy control information 245. On the contrary, since it is necessary to suppress copy processing when the desired value is smaller, a smaller copy speed is set as the copy control information 245.

[0070] In the above example [ 6th ], even if an input/output request arrival rate changes during disk restoration processing, the automatic regulation of the copy processing can be carried out, and the aim engine performance can be maintained without a break in of an operator. In addition, since the priority of copy processing is changeable even if it changes the gap which sets up the right 243 of copy processing activation by the dispatch control section 234 outside it changes copy speed, the same effect as the above is acquired. Therefore, the copy controller 238 may be made to make an automatic change of the time interval which sets up the right 243 of copy processing activation.

[0071] - The 7th example-7th example is automatic setting / example which carries out an automatic regulation about copy speed using the average engine-performance maintenance time amount THP. Drawing 12 is the block diagram showing disk array system 100C of the 7th example. First, the average engine-performance maintenance time amount THP over each copy speed is calculated like the 4th example, the copy speed from which the average engine-performance maintenance time amount THP becomes max is set automatically, and copy processing is started.

[0072] In the monitor section 235, while always measuring the response time, the elapsed time from the number of trucks and copy initiation which ended the copy is measured, and it is set as the monitor result 247. The current condition under disk restoration processing activation is classified into an engine-performance maintenance condition or a performance degradation condition according to the index presumption section 236 by comparing with the aim response time (desired value 250) the

response time read from the monitor result 247. Moreover, the disk release time T is re-calculated based on the monitor result 247. Next, the average engine-performance maintenance time amount THP is re-calculated using the re-calculated disk release time T. And the average engine-performance maintenance time amount THP is set as the presumed result 249.

[0073] When the numeric value of the average engine-performance maintenance time amount THP is getting worse rather than last time with reference to the presumed result 249 the suitable period, directions are taken out with the copy controller 238 to the index presumption section 236 so that the average engine-performance maintenance time amount THP when changing copy speed may be re-presumed. In the index presumption section 236, the average engine-performance maintenance time amount THP over each copy speed is re-presumed anew, and it is set as the presumed result 249. In the copy controller 238, the copy speed from which the average engine-performance maintenance time amount THP becomes max is found based on this presumed result 249, and it is set as the copy control information 245.

[0074] In the above example [ 7th ], even if an input/output request arrival rate changes during disk restoration processing, the automatic regulation of the copy speed can be carried out without a break in of an operator so that the balance of reliability and the on-line-processing engine performance may become the optimal.

[0075] In addition, the 7th example may be changed into either of the degrees according to the capacity of a system.

\*\* When the numeric value of the average engine-performance maintenance time amount THP is getting worse with reference to the presumed result 249 the suitable period from the average engine-performance maintenance time amount THP calculated at the time of copy processing initiation, the copy controller 238 takes out directions to the index presumption section 236 so that the average engine-performance maintenance time amount THP when changing copy speed may be re-presumed.

\*\* By comparing with the aim response time (desired value 250) the response time read from the monitor result 247, the index presumption section 236 classifies the current condition under disk restoration processing activation into an engine-performance maintenance condition or a performance degradation condition, and sets it as the presumed result 249. When it changes into a performance degradation condition with reference to the presumed result 249 a suitable period, the copy controller 238 takes out directions to the index presumption section 236 so that the average engine-performance maintenance time amount THP when changing copy speed may be re-presumed.

\*\* The index presumption section 236 re-calculates the disk release time T with reference to the monitor result 247 a suitable period. Next, the average engine-performance maintenance time amount THP over each copy speed is re-calculated using the re-calculated disk release time T. And the average engine-performance maintenance time amount THP is set as the presumed result 249.

[0076] - An 8th example-service processor can be connected to a host computer 1, and the configuration with which the disk restoration initiation directions from an operator reach a disk controller 2 via a service processor from a host computer 1, then an operator can direct disk restoration initiation from a host computer 1. Moreover, an operator can direct disk restoration initiation from a host computer 1 also as a configuration which prepared the function of a service processor in the host computer 1.

[0077] - Although the operator is directing disk restoration initiation in the 9th example-1st example -, and the 8th example after fixing and exchanging the broken disk drive, start disk restoration processing automatically in the 9th example, without waiting for repair and exchange of a disk drive. Drawing 13 is the block diagram showing disk array system 100D of the 9th example. Disk array system 100D has five sets of disk drives 31-35. The disk drive 35 of these is a disk drive of a spare, and while the system is working normally, it is not accessed.

[0078] Moreover, the disk failure detection section 239 is formed in the disk controller 2. It writes the broken disk drive number in fault information 251 while it will set up the copy initiation directions 244, if this disk failure detection section 239 has the function to detect the disk failure under radial transfer activation and a disk failure is detected.

[0079] In the copy controller 238, when the copy initiation directions 244 are set up, with reference to

the aim response time and the monitor result 247 which are beforehand registered into desired value 250, the copy control information 245 is set up like the 6th example. In the copy processing section 233, based on fault information 251 and the copy control information 245, data is read from three sets of the normal disks of the disk drives 31-34, and the restored data is written in to the spare drive 35.

[0080] In the above example [ 9th ], even if one set of a disk drive breaks down, disk restoration processing is started automatically, without waiting for repair and exchange of the disk drive, and disk restoration processing can be performed, protecting the aim engine performance set up beforehand. In addition, instead of the response time, the average engine-performance maintenance time amount THP may be used like the 7th example, and disk restoration processing may be performed.

[0081]

[Effect of the Invention] According to the disk array system of this invention, in case disk restoration processing is performed, the index for determining to copy processings [ how many ] priority should be given is calculated, an operator can be shown or optimal control can be performed based on it. Therefore, suitable disk restoration processing can be performed now from a viewpoint of both reliability and the on-line-processing engine performance.

---

[Translation done.]

(19) 日本国特許庁 (JP)

(12) 公開特許公報 (A)

(11)特許出願公開番号

特開平6-250795

(43)公開日 平成6年(1994)9月9日

(51) Int.Cl.<sup>5</sup>  
G 06 F 3/06

識別記号 庁内整理番号  
304 N 7165-5B  
301 Z 7165-5B

F I

技術表示箇所

審査請求 未請求 請求項の数11 OL (全 21 頁)

(21)出願番号 特願平5-35233

(22)出願日 平成5年(1993)2月24日

(71)出願人 000005108

株式会社日立製作所

東京都千代田区神田駿河台四丁目 6 番地

赤津 雅晴

神奈川県川崎市麻生区王禅寺1099番地 株式会社日本製作所(ニッポン開発研究所)

立日社芸芸

村田　省介  
神奈川県川崎市麻生区王禅寺1099番地 株

式会社日立

栗原 謙三  
神奈川県川崎市麻生区王禅寺1099番地 株

式会社日立製作所システム

最終百に綴ぐ

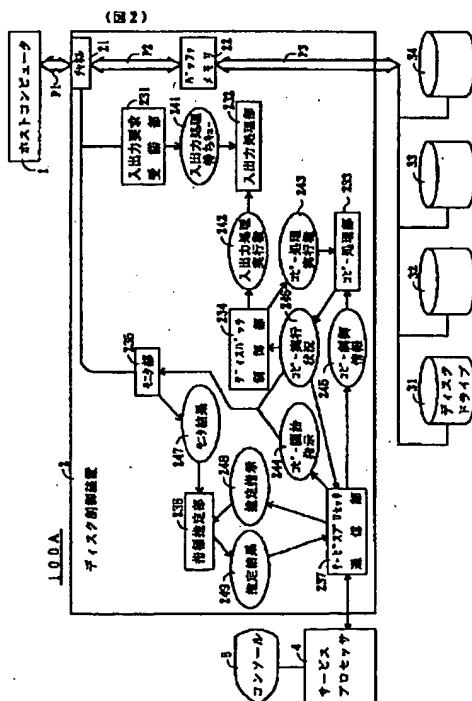
(54) 【発明の名称】 ディスクアレイシステム

(57) 【要約】

【目的】 ディスク復旧処理を行う際に、コピー処理をどの程度優先すべきかを決定するための指標を操作員に提示したり、それに基づいて最適な制御を行うディスクアレイシステムを提供する。

【構成】 モニタ部235にて入出力要求到着率、平均応答時間、ディスク復旧時間を計測して、モニタ結果236として記録する。ディスク復旧開始時あるいは実行中に、指標推定部236で、現時点の入出力要求到着率と前記モニタ結果236を用いて、ディスク復旧時間、平均応答時間、平均性能維持時間を推定し、コピー速度を指定する際の参考として操作員に提示したり、それに基づいて自動的に最適な制御を行ってコピー処理を実行する。

【効果】 信頼性とオンライン処理性能の両方の観点から適切なディスク復旧処理を行うことが出来る。



1

2

## 【特許請求の範囲】

【請求項1】 ホストコンピュータとディスク制御装置と複数のディスクドライブとから構成され、データを複数のディスクドライブに分散して記録し、1台以上のディスクドライブが故障したとき、故障したディスクの保持していたデータを正常なディスクドライブ内のデータをもとに復元して、修理・交換したディスクドライブまたはスペアのディスクドライブに書き込むコピー処理を、操作員に指定されたコピー速度で実行することによって、オンライン処理を継続しながら、前記故障に対する復旧処理を行なうディスクアレイシステムにおいて、ホストコンピュータからの入出力要求到着率と、コピー処理完了までのディスク復旧時間と、コピー速度と、ディスク復旧処理中の平均入出力要求到着率とを計測し、それらをモニタ結果として記録するモニタ手段と、前記モニタ結果をもとに、現時点での入出力要求到着率が維持されたと仮定したときのコピー速度とディスク復旧時間の関係を求め、それを推定結果として出力する指標推定手段と、

前記推定結果をコピー速度指定用指標として表示し、操作員にコピー速度を指定させるサービス手段とを具備したことを特徴とするディスクアレイシステム。

【請求項2】 ホストコンピュータとディスク制御装置と複数のディスクドライブとから構成され、データを複数のディスクドライブに分散して記録し、1台以上のディスクドライブが故障したとき、故障したディスクの保持していたデータを正常なディスクドライブ内のデータをもとに復元して、修理・交換したディスクドライブまたはスペアのディスクドライブに書き込むコピー処理を、操作員に指定されたコピー速度で実行することによって、オンライン処理を継続しながら、前記故障に対する復旧処理を行なうディスクアレイシステムにおいて、ホストコンピュータからの入出力要求到着率と、コピー処理完了までのディスク復旧時間と、コピー速度と、ディスク復旧処理中の平均入出力要求到着率とを計測し、それらをモニタ結果として記録するモニタ手段と、

操作員に指定されたコピー速度および前記モニタ結果をもとに、現時点での入出力要求到着率が維持されたと仮定したときのディスク復旧時間を求め、それを推定結果として出力する指標推定手段と、

前記推定結果を表示するサービス手段とを具備したことを特徴とするディスクアレイシステム。

【請求項3】 ホストコンピュータとディスク制御装置と複数のディスクドライブとから構成され、データを複数のディスクドライブに分散して記録し、1台以上のディスクドライブが故障したとき、故障したディスクの保持していたデータを正常なディスクドライブ内のデータをもとに復元して、修理・交換したディスクドライブまたはスペアのディスクドライブに書き込むコピー処理を、操作員に指定されたコピー速度で実行することによ

10

20

30

40

50

って、オンライン処理を継続しながら、前記故障に対する復旧処理を行なうディスクアレイシステムにおいて、ホストコンピュータからの入出力要求到着率と、コピー処理完了までのディスク復旧時間と、コピー速度と、ディスク復旧処理中の平均入出力要求到着率とを計測し、それらをモニタ結果として記録するモニタ手段と、

前記モニタ結果をもとに、現時点での入出力要求到着率が維持されたと仮定したときのコピー速度とディスク復旧時間の関係を求め、さらに、そのディスク復旧時間と1台のディスクドライブの故障率と修理交換率とをもとに複数台のディスク装置の故障によりデータ復元が不可能となるまでの平均時間MTTDLを計算し、それを推定結果として出力する指標推定手段と、

前記推定結果をコピー速度指定用指標として表示し、操作員にコピー速度を指定させるサービス手段とを具備したことを特徴とするディスクアレイシステム。

【請求項4】 ホストコンピュータとディスク制御装置と複数のディスクドライブとから構成され、データを複数のディスクドライブに分散して記録し、1台以上のディスクドライブが故障したとき、故障したディスクの保持していたデータを正常なディスクドライブ内のデータをもとに復元して、修理・交換したディスクドライブまたはスペアのディスクドライブに書き込むコピー処理を、操作員に指定されたコピー速度で実行することによって、オンライン処理を継続しながら、前記故障に対する復旧処理を行なうディスクアレイシステムにおいて、ホストコンピュータからの入出力要求到着率と、ディスク復旧処理中のホストコンピュータからの入出力要求の処理完了までの平均応答時間と、コピー速度と、ディスク復旧処理中の平均入出力要求到着率とを計測し、それらをモニタ結果として記録するモニタ手段と、

前記モニタ結果をもとに、現時点での入出力要求到着率が維持されたと仮定したときのコピー速度と平均応答時間の関係を求め、それを推定結果として出力する指標推定手段と、

前記推定結果をコピー速度指定用指標として表示し、操作員にコピー速度を指定させるサービス手段とを具備したことを特徴とするディスクアレイシステム。

【請求項5】 ホストコンピュータとディスク制御装置と複数のディスクドライブとから構成され、データを複数のディスクドライブに分散して記録し、1台以上のディスクドライブが故障したとき、故障したディスクの保持していたデータを正常なディスクドライブ内のデータをもとに復元して、修理・交換したディスクドライブまたはスペアのディスクドライブに書き込むコピー処理を、操作員に指定されたコピー速度で実行することによって、オンライン処理を継続しながら、前記故障に対する復旧処理を行なうディスクアレイシステムにおいて、ホストコンピュータからの入出力要求到着率と、ホストコンピュータからの入出力要求の処理完了までの応答時

間と、コピー処理完了までのディスク復旧時間と、コピー速度と、ディスク復旧処理中の平均入出力要求到着率とを計測し、それらをモニタ結果として記録するモニタ手段と、

前記モニタ結果をもとに、現時点での入出力要求到着率が維持されたと仮定したときのコピー速度とディスク復旧時間の関係を求め、また、操作員に指定された目標応答時間と前記モニタ結果をもとに、現時点での入出力要求到着率が維持されたと仮定したときの各コピー速度に対する応答時間を推定し、その応答時間が目標応答時間より小さい場合は性能維持状態とし、大きい場合は性能劣化状態とする状態分類を行い、前記ディスク復旧時間と1台のディスクドライブの故障率と修理交換率と前記状態分類とをもとに複数台のディスクドライブの故障によりデータ復元が不可能となるまでの間に前記性能維持状態で機能している平均性能維持時間を計算して、それを推定結果として出力する指標推定手段と、

前記推定結果をコピー速度指定用指標として表示し、操作員にコピー速度を指定させるサービス手段とを具備したことを特徴とするディスクアレイシステム。

【請求項6】 請求項1から請求項5に記載のディスクアレイシステムにおいて、

モニタ手段は、コピー処理実行状況をチェックし、モニタ結果に記録し、

指標推定手段は、コピー処理実行中に、前記モニタ結果をもとにディスク復旧に要する残り時間を推定し、推定結果に記録し、

サービス手段は、コピー処理実行中に、前記推定結果をコピー速度変更指示用指標として表示し、操作員にコピー速度を変更指示させ、

さらに、コピー処理操作員からコピー速度変更指示があったときにコピー速度を変更してコピー処理を行うコピー処理手段を具備したことを特徴とするディスクアレイシステム。

【請求項7】 請求項1から請求項5に記載のディスクアレイシステムにおいて、

モニタ手段は、コピー処理実行状況をチェックし、モニタ結果に記録し、

サービス手段は、操作員から目標応答時間を受け取り、さらに、コピー処理実行中に、前記モニタ結果と前記目標応答時間をもとに、平均応答時間が目標応答時間より

小さい場合にはコピー速度を大きめに変更するか又はコピー実行の頻度を高めに変更し、平均応答時間が目標応答時間より大きい場合にはコピー速度を小さめに変更するか又はコピー実行の頻度を低めに変更するコピー調整手段と、

そのコピー調整手段によるコピー条件の変更があったときにコピー条件を変更してコピー処理を行うコピー処理手段とを具備したことを特徴とするディスクアレイシステム。

【請求項8】 請求項7に記載のディスクアレイシステムにおいて、スペアのディスクドライブを有し且つ目標応答速度およびコピー開始時のコピー速度が予め与えられている場合、操作員の指示を待たずに、各手段がコピー処理のための動作を開始することを特徴とするディスクアレイシステム。

【請求項9】 ホストコンピュータとディスク制御装置と複数のディスクドライブとから構成され、データを複数のディスクドライブに分散して記録し、1台以上のディスクドライブが故障したとき、故障したディスクの保持していたデータを正常なディスクドライブ内のデータをもとに復元して、修理・交換したディスクドライブまたはスペアのディスクドライブに書き込むコピー処理を、所定のコピー速度で実行することによって、オンライン処理を継続しながら、前記故障に対する復旧処を行なうディスクアレイシステムにおいて、

ホストコンピュータからの入出力要求到着率と、ホストコンピュータからの入出力要求の処理完了までの応答時間と、コピー処理完了までのディスク復旧時間と、コピー速度と、ディスク復旧処理中の平均入出力要求到着率とを計測し、それらをモニタ結果として記録するモニタ手段と、

前記モニタ結果をもとに、現時点での入出力要求到着率が維持されたと仮定したときのコピー速度とディスク復旧時間の関係を求め、また、操作員に指定された目標応答時間と前記モニタ結果をもとに、現時点での入出力要求到着率が維持されたと仮定したときの各コピー速度に対する応答時間を推定し、その平均応答時間が目標応答時間より小さい場合は性能維持状態とし、大きい場合は性能劣化状態とする状態分類を行い、前記ディスク復旧時間と1台のディスクドライブの故障率と修理交換率と前記状態分類とをもとに複数台のディスクドライブの故障によりデータ復元が不可能となるまでの間に前記性能維持状態で機能している平均性能維持時間を計算して、それを推定結果として出力する指標推定手段と、

前記推定結果を参照し、平均性能維持時間が最大となるコピー速度を設定するコピー調整手段と、そのコピー調整手段により設定されたコピー速度によりコピー処理を開始するコピー処理手段とを具備したことを特徴とするディスクアレイシステム。

【請求項10】 請求項9に記載のディスクアレイシステムにおいて、

モニタ手段は、応答時間を計測すると共に、コピーを終了したトラック数とコピー開始からの経過時間を計測して、モニタ結果に設定し、

指標推定手段は、コピー処理実行中に、操作員に指定された目標応答時間とモニタ結果をもとに、各コピー速度に対する平均性能維持時間を再推定し、それを推定結果に設定し、

50 コピー調整手段は、適当な周期で前記推定結果を参照

し、平均性能維持時間が最大となるコピー速度に変更し、

コピー手段は、前記コピー調整手段によりコピー速度の変更があったときにそのコピー速度に変更してコピー処理を行うことを特徴とするディスクアレイシステム。

【請求項11】 請求項9または請求項10に記載のディスクアレイシステムにおいて、スペアのディスクドライブを有し且つ目標応答速度が予め与えられている場合、操作員の指示を待たずに、各手段がコピー処理のための動作を開始することを特徴とするディスクアレイシステム。 10

【発明の詳細な説明】

【0001】

【産業上の利用分野】 本発明は、ディスクアレイシステムに関し、さらに詳しくは、ディスクドライブが故障したときのディスク復旧処理を好適に行なうことが出来るディスクアレイシステムに関する。

【0002】

【従来の技術】 ディスクアレイシステムは、複数のディスクドライブに分散してデータを記録し、1台以上のディスクドライブが故障したとき、その故障したディスクドライブの保持していたデータを、正常なディスクドライブ内のデータをもとに復元し、修理・交換したディスクドライブまたはスペアのディスクドライブに書き込むコピー処理を、数トラックずつ行なうことによって、オンライン処理を継続しながら、前記故障に対する復旧処理を行えるようにしたシステムである。このようなディスクアレイシステムは、例えば、「普及へ向け、製品化進むディスク・アレイ(日経コンピュータ1992.3.9;87~98ページ)」に記載されている。

【0003】

【発明が解決しようとする課題】 ディスクアレイシステムにおいて、コピー処理の優先度を上げていくと、一部のディスクドライブが故障している状態に留まっている時間が短くなるので、多重障害によってデータ復元が不可能になってしまふ状態に陥る確率は低下する。すなわち、信頼性は向上する。その反面、その分だけオンライン処理が疎かになるため、オンライン処理性能は低下する。すなわち、ディスクアレイシステムでは、オンライン処理とコピー処理とを並列に行なうが、コピー処理をオンライン処理に比べてどの程度優先して行なうかを決定するのが難しい。しかし、従来、コピー処理の優先度に関する技術は開示されていない。

【0004】 本発明は、このような事情に鑑みてなされたもので、その目的とするところは、ディスク復旧処理を行う際に、コピー処理をどの程度優先すべきかを決定するための指標を計算して、操作員に提示したり、それに基づいて最適な制御を行い、信頼性とオンライン処理性能の両方の観点から適切なディスク復旧処理を行うディスクアレイシステムを提供することにある。

【0005】

【課題を解決するための手段】 第1の観点では、本発明は、ホストコンピュータからの入出力要求到着率と、コピー処理完了までのディスク復旧時間と、コピー速度と、ディスク復旧処理中の平均入出力要求到着率とを計測し、それらをモニタ結果として記録するモニタ手段と、前記モニタ結果をもとに、現時点での入出力要求到着率が維持されたと仮定したときのコピー速度とディスク復旧時間の関係を求め、それを推定結果として出力する指標推定手段と、前記推定結果をコピー速度指定用指標として表示し、操作員にコピー速度を指定させるサービス手段とを具備したことを特徴とするディスクアレイシステムを提供する。

【0006】 第2の観点では、本発明は、ホストコンピュータからの入出力要求到着率と、コピー処理完了までのディスク復旧時間と、コピー速度と、ディスク復旧処理中の平均入出力要求到着率とを計測し、それらをモニタ結果として記録するモニタ手段と、操作員に指定されたコピー速度および前記モニタ結果をもとに、現時点での入出力要求到着率が維持されたと仮定したときのディスク復旧時間を求め、それを推定結果として出力する指標推定手段と、前記推定結果を表示するサービス手段とを具備したことを特徴とするディスクアレイシステムを提供する。

【0007】 第3の観点では、本発明は、ホストコンピュータからの入出力要求到着率と、コピー処理完了までのディスク復旧時間と、コピー速度と、ディスク復旧処理中の平均入出力要求到着率とを計測し、それらをモニタ結果として記録するモニタ手段と、前記モニタ結果をもとに、現時点での入出力要求到着率が維持されたと仮定したときのコピー速度とディスク復旧時間の関係を求め、さらに、そのディスク復旧時間と1台のディスクドライブの故障率と修理交換率とをもとに複数台のディスク装置の故障によりデータ復元が不可能となるまでの平均時間MTTDLを計算し、それを推定結果として出力する指標推定手段と、前記推定結果をコピー速度指定用指標として表示し、操作員にコピー速度を指定させるサービス手段とを具備したことを特徴とするディスクアレイシステムを提供する。

【0008】 第4の観点では、本発明は、ホストコンピュータからの入出力要求到着率と、ディスク復旧処理中のホストコンピュータからの入出力要求の処理完了までの平均応答時間と、コピー速度と、ディスク復旧処理中の平均入出力要求到着率とを計測し、それらをモニタ結果として記録するモニタ手段と、前記モニタ結果をもとに、現時点での入出力要求到着率が維持されたと仮定したときのコピー速度と平均応答時間の関係を求め、それを推定結果として出力する指標推定手段と、前記推定結果をコピー速度指定用指標として表示し、操作員にコピー速度を指定させるサービス手段とを具備したことを特徴とするディスクアレイシステムを提供する。

7 特徴とするディスクアレイシステムを提供する。

【0009】第5の観点では、本発明は、ホストコンピュータからの入出力要求到着率と、ホストコンピュータからの入出力要求の処理完了までの応答時間と、コピー処理完了までのディスク復旧時間と、コピー速度と、ディスク復旧処理中の平均入出力要求到着率とを計測し、それらをモニタ結果として記録するモニタ手段と、前記モニタ結果をもとに、現時点での入出力要求到着率が維持されたと仮定したときのコピー速度とディスク復旧時間の関係を求め、また、操作員に指定された目標応答時間と前記モニタ結果をもとに、現時点での入出力要求到着率が維持されたと仮定したときの各コピー速度に対する応答時間を推定し、その応答時間が目標応答時間より小さい場合は性能維持状態とし、大きい場合は性能劣化状態とする状態分類を行い、前記ディスク復旧時間と1台のディスクドライブの故障率と修理交換率と前記状態分類とをもとに複数台のディスクドライブの故障によりデータ復元が不可能となるまでの間に前記性能維持状態で機能している平均性能維持時間を計算して、それを推定結果として出力する指標推定手段と、前記推定結果をコピー速度指定用指標として表示し、操作員にコピー速度を指定させるサービス手段とを備したことを特徴とするディスクアレイシステムを提供する。

【0010】第6の観点では、本発明は、上記構成のディスクアレイシステムにおいて、モニタ手段は、コピー処理実行状況をチェックし、モニタ結果に記録し、指標推定手段は、コピー処理実行中に、前記モニタ結果をもとにディスク復旧に要する残り時間を推定し、推定結果に記録し、サービス手段は、コピー処理実行中に、前記推定結果をコピー速度変更指示用指標として表示し、操作員にコピー速度を変更指示させ、さらに、コピー処理操作員からコピー速度変更指示があったときにコピー速度を変更してコピー処理を行うコピー処理手段を備したことを特徴とするディスクアレイシステムを提供する。

【0011】第7の観点では、本発明は、上記構成のディスクアレイシステムにおいて、モニタ手段は、コピー処理実行状況をチェックし、モニタ結果に記録し、サービス手段は、操作員から目標応答時間を受け取り、さらに、コピー処理実行中に、前記モニタ結果と前記目標応答時間をもとに、平均応答時間が目標応答時間より小さい場合にはコピー速度を大きめに変更するか又はコピー実行の頻度を高めに変更し、平均応答時間が目標応答時間より大きい場合にはコピー速度を小さめに変更するか又はコピー実行の頻度を低めに変更するコピー調整手段と、そのコピー調整手段によるコピー条件の変更があったときにコピー条件を変更してコピー処理を行うコピー処理手段とを備したことを特徴とするディスクアレイシステムを提供する。

【0012】第8の観点では、本発明は、上記第7の観

50

点によるディスクアレイシステムにおいて、スペアのディスクドライブを有し且つ目標応答時間およびコピー開始時のコピー速度が予め与えられている場合、操作員の指示を待たずに、各手段がコピー処理のための動作を開始することを特徴とするディスクアレイシステムを提供する。

【0013】第9の観点では、本発明は、ホストコンピュータからの入出力要求到着率と、ホストコンピュータからの入出力要求の処理完了までの応答時間と、コピー処理完了までのディスク復旧時間と、コピー速度と、ディスク復旧処理中の平均入出力要求到着率とを計測し、それらをモニタ結果として記録するモニタ手段と、前記モニタ結果をもとに、現時点での入出力要求到着率が維持されたと仮定したときのコピー速度とディスク復旧時間の関係を求め、また、操作員に指定された目標応答時間と前記モニタ結果をもとに、現時点での入出力要求到着率が維持されたと仮定したときの各コピー速度に対する応答時間を推定し、その平均応答時間が目標応答時間より小さい場合は性能維持状態とし、大きい場合は性能劣化状態とする状態分類を行い、前記ディスク復旧時間と1台のディスクドライブの故障率と修理交換率と前記状態分類とをもとに複数台のディスクドライブの故障によりデータ復元が不可能となるまでの間に前記性能維持状態で機能している平均性能維持時間を計算して、それを推定結果として出力する指標推定手段と、前記推定結果を参考し、平均性能維持時間が最大となるコピー速度を設定するコピー調整手段と、そのコピー調整手段により設定されたコピー速度によりコピー処理を開始するコピー処理手段とを備したことを特徴とするディスクアレイシステムを提供する。

【0014】第10の観点では、本発明は、上記第9の観点によるディスクアレイシステムにおいて、モニタ手段は、応答時間を計測すると共に、コピーを終了したトラック数とコピー開始からの経過時間を計測して、モニタ結果に設定し、指標推定手段は、コピー処理実行中に、操作員に指定された目標応答時間とモニタ結果をもとに、各コピー速度に対する平均性能維持時間を再推定し、それを推定結果に設定し、コピー調整手段は、適当な周期で前記推定結果を参照し、平均性能維持時間が最大となるコピー速度に変更し、コピー手段は、前記コピー調整手段によりコピー速度の変更があったときにそのコピー速度に変更してコピー処理を行うことを特徴とするディスクアレイシステムを提供する。

【0015】第11の観点では、本発明は、上記第9の観点または第10の観点によるディスクアレイシステムにおいて、スペアのディスクドライブを有し且つ目標応答速度が予め与えられている場合、操作員の指示を待たずに、各手段がコピー処理のための動作を開始することを特徴とするディスクアレイシステムを提供する。

【0016】

【作用】上記第1の観点による本発明のディスクアレイシステムでは、コピー処理をどの程度優先すべきかを決定するための指標として、現時点での入出力要求到着率が維持されたと仮定したときのコピー速度とディスク復旧時間の関係を、操作員に提示する。そこで、操作員は、信頼性とオンライン処理性能の両方の観点から適切なディスク復旧時間となるようなコピー速度を指定できる。

【0017】上記第2の観点による本発明のディスクアレイシステムでは、操作員に指定されたコピー速度でコピー処理し且つ現時点での入出力要求到着率が維持されたと仮定したときのディスク復旧時間を、操作員に提示する。そこで、操作員は、信頼性とオンライン処理性能の両方の観点から適切なディスク復旧時間となるようなコピー速度を指定したか否かを確認できる。

【0018】上記第3の観点による本発明のディスクアレイシステムでは、コピー処理をどの程度優先すべきかを決定するための指標として、データ復元が不可能となるまでの平均時間MTTDLを、操作員に提示する。そこで、操作員は、信頼性とオンライン処理性能の両方の観点から適切なディスク復旧時間となるようなコピー速度を指定できる。

【0019】上記第4の観点による本発明のディスクアレイシステムでは、コピー処理をどの程度優先すべきかを決定するための指標として、現時点での入出力要求到着率が維持されたと仮定したときのコピー速度と平均応答時間の関係を、操作員に提示する。そこで、操作員は、信頼性とオンライン処理性能の両方の観点から適切な平均応答時間となるようなコピー速度を指定できる。

【0020】上記第5の観点による本発明のディスクアレイシステムでは、各コピー速度に対する応答時間が目標応答時間より小さい場合を性能維持状態と呼び、データ復元が不可能となるまでの間に前記性能維持状態で機能している総時間を平均性能維持時間と呼ぶとき、その平均性能維持時間を、コピー処理をどの程度優先すべきかを決定するための指標として、操作員に提示する。そこで、操作員は、信頼性とオンライン処理性能の両方の観点から適切な平均性能維持時間となるようなコピー速度を指定できる。

【0021】上記第6の観点による本発明のディスクアレイシステムでは、コピー処理実行中に、ディスク復旧に要する残り時間を推定し、それを操作員に提示し、コピー速度を変更指示させ、コピー処理操作員が変更したコピー速度でコピー処理を継続する。そこで、コピー処理実行中に入出力要求到着率が変動した場合などに対応できるようになる。

【0022】上記第7の観点による本発明のディスクアレイシステムでは、コピー処理実行中に、平均応答時間が目標応答時間に一致するように、コピー速度またはコピー実行の頻度を自動調整する。そこで、コピー処理実

行中に入出力要求到着率が変動した場合などに対応できるようになる。

【0023】上記第8の観点による本発明のディスクアレイシステムでは、スペアのディスクドライブを有しているとき、操作員の指示を待たずに、予め与えられているコピー開始時のコピー速度で、各手段がコピー処理のための動作を開始する。そこで、操作員の指示が不要となり、保守を自動化できる。

【0024】上記第9の観点による本発明のディスクアレイシステムでは、平均性能維持時間が最大になるように、コピー速度を自動設定し、コピー処理を開始する。そこで、操作員の指示が不要となる。上記第10の観点による本発明のディスクアレイシステムでは、コピー処理実行中に、平均性能維持時間が最大になるように、コピー速度を自動調整する。そこで、コピー処理実行中に入出力要求到着率が変動した場合などに対応できるようになる。上記第11の観点による本発明のディスクアレイシステムでは、スペアのディスクドライブを有しているとき、操作員の指示を待たずに、予め与えられているコピー開始時のコピー速度で、各手段がコピー処理のための動作を開始する。そこで、操作員の指示が不要となり、保守を自動化できる。

#### 【0025】

【実施例】以下、本発明の実施例を図面により詳細に説明する。なお、これにより本発明が限定されるものではない。図1は、後述する第1実施例から第9実施例にかかるディスクアレイシステムに共通するハードウエア構成のブロック図である。このディスクアレイシステム100は、ホストコンピュータ1、ディスク制御装置2、ディスクドライブ31～34、サービスプロセッサ4およびコンソール5から構成されている。本実施例では、4台のディスクドライブ31～34の内の3台にデータを分散して書き込み、残りの1台にエラー訂正コードを書き込むものとする。ただし、本発明においてディスクドライブの台数に制約はなく、m台のディスクドライブを用意して、そのうちのn台をエラー訂正ドライブとしてよい(m, nはm>n≥1を満たす任意の自然数)。ディスク制御装置2は、チャネル21、バッファメモリ22、制御プロセッサ25, 26、ローカルメモリ27, 28および共通メモリ29から構成されている。第1実施例から第9実施例は、ソフトウェア構成するわち機能がそれぞれ異なっている。次に、第1実施例から第9実施例を順に説明する。

#### 【0026】-第1実施例-

図2は、本発明の第1実施例にかかるディスクアレイシステム100Aの構成図である。ディスク制御装置2は、チャネル21とバッファメモリ22を有する。ホストコンピュータ1とチャネル21の間、チャネル21とバッファメモリ22の間、バッファメモリ22とディスクドライブ31～34の間に、それぞれデータ転送路P

1, P2, P3がある。

【0027】ディスク制御装置2内にある四角で囲まれた各部231～237は、制御プロセッサ25, 26のマイクロプログラムにより実現される機能を表している。なお、各機能毎に1台の制御プロセッサを割り当てても良いし、1台の制御プロセッサだけに全ての機能を持たせてもよい。また、ディスク制御装置2内にある梢円で囲まれた各情報241～249は、上記各機能間で受け渡しを行う情報である。これらの情報のうち異なる制御プロセッサ25, 26に割り当てられている機能間で受け渡しを行う情報は、各制御プロセッサ25, 26が参照・更新できる共通メモリ29内に置く。また、同一制御プロセッサに割り当てられている機能間で受け渡しを行う情報は、その制御プロセッサのローカルメモリ27または28に置く。

【0028】入出力要求受諾部231は、ホストコンピュータ1が、チャネル21を介して、ディスクドライブ31～34に対する入出力要求（データ読み出し命令／データ書き込み命令）を発行したとき、それを検出して、入出力処理待ちキュー241に登録する。

【0029】入出力処理部232は、ディスパッチ制御部234から入出力処理実行権242が与えられたとき、入出力処理待ちキュー241の先頭に登録されている入出力要求を処理する。入出力要求がデータ読み出し命令の場合は、データを格納している3台のディスクドライブが正常であれば、その3台のディスクドライブをアクセスして、データを読み込む。そして、1台が故障していれば、残り3台からデータとエラー訂正用コードとを読み込み、故障しているディスクドライブに格納されているデータを復元したのち、パッファメモリ22を経由して、ホストコンピュータ1に転送する。入出力要求がデータ書き込み命令の場合は、データを3つに分割するとともに、エラー訂正用コードを作成して、パッファメモリ22を経由して、それぞれを正常なディスクドライブに書き込む。

【0030】コピー処理部233は、4台のディスクドライブの内の1台が故障したときに、それを修理・交換した後、コピー処理を行う。図3に、コピー処理のフロー図を示す。まず、サービスプロセッサ通信部237からコピー開始指示244を受けるまでループする（ステップ1000）。コピー開始指示244を受けると、復旧対象のディスクドライブ番号と1回のコピー処理で扱うトラック数（以下、コピー速度と呼ぶ。）から成るコピー制御情報245を読み込む（ステップ1010）。次に、コピー処理を行うトラック番号を初期化し（ステップ1020）、ディスパッチ制御部234から、コピー処理実行権243が与えられるまでループする（ステップ1030）。

【0031】コピー処理実行権243が与えられると、これからコピー処理を行う先頭のトラック番号を読み出

す（ステップ1040）。次に、正常な3台のディスクドライブからデータを読み出して、パッファメモリ22に格納する（ステップ1050）。データを読み出すトラックは、上記トラック番号から始まって上記コピー速度で指示された数のトラックである。

【0032】次に、パッファメモリ22に読み出したデータから、復旧対象のディスクドライブのデータを復元して、パッファメモリ22に格納する（ステップ1060）。そして、復元したデータをパッファメモリ22から復旧対象のディスクドライブに転送する（ステップ1070）。

【0033】次に、次回のコピー処理対象の先頭トラック番号を更新する（ステップ1080）。最後に、ディスク復旧処理が完了したか否かを判定し（ステップ1090）、完了していないければ、前記ステップ1030に戻る。完了したならば、コピー実行状況246にコピー処理完了を設定し（ステップ1100）、前記ステップ1000に戻る。なお、入出力処理部232は、入出力処理実行権242を検出した後、その入出力処理実行権242をリセットする。また、コピー処理部233は、コピー処理実行権243を検出した後、そのコピー処理実行権243をリセットする。

【0034】ディスパッチ制御部234は、適当なスケジュール処理で、入出力処理部232に入出力処理実行権242を与える。また、コピー処理部233にコピー処理実行権を与える。図4に、スケジュール処理のフロー図を示す。このスケジュール処理は、通常は入出力処理実行権242を設定し、一定の周期ごとにコピー処理実行権を設定するものである。まず、サービスプロセッサ通信部237からコピー開始指示244を受けるまでの間は、入出力処理実行権242を設定する（ステップ2000, 2010）。

【0035】サービスプロセッサ通信部237からコピー開始指示244を受けると、タイマを0リセットする（ステップ2020）。そして、タイマをチェックして、所定の設定時間（例えば50ms）を経過していないければ、入出力処理実行権242を設定する（ステップ2030, 2040）。所定の設定時間が経過したならば、コピー処理実行権243を設定する（ステップ2050）。

【0036】そして、コピー実行状況246に、コピー処理完了が設定されていなければ前記ステップ2020に戻り、設定されていなければ前記ステップ2000に戻る（ステップ2060）。なお、各処理実行権242, 243の設定処理（ステップ2010, 2040, 2050）では、前回に設定した処理実行権がリセットされるまで、その場でループする。

【0037】モニタ部235は、チャネル21に入っているホストコンピュータ1からの入出力要求を常時チェックして、入出力要求到着率を計測する。そして、その

入出力要求到着率をモニタ結果247に記録する。また、コピー開始指示244が設定されてからコピー実行状況246にコピー処理完了が設定されるまでのディスク復旧時間と、コピー速度と、ディスク復旧処理中の平均入出力要求到着率をモニタ結果247に記録する。これらディスク復旧時間、コピー速度、平均入出力要求到着率は、ディスク復旧処理が行われる度にモニタされ、統計処理されて、図5に示すような関係が得られる。

【0038】指標推定部236は、サービスプロセッサ通信部237より推定指示248を受けたとき、モニタ結果247からその時点での入出力要求到着率を読み出し、その入出力要求到着率が維持されたと仮定したときのコピー速度とディスク復旧時間の関係を図5の関係から求めて、その結果を推定結果249に設定する。ある入出力要求到着率におけるコピー速度とディスク復旧時間の関係は、図6に示すようになる。

【0039】サービスプロセッサ通信部237は、以下の3つの機能を持つ。第一に、サービスプロセッサ4よりディスク復旧処理開始を指示されたとき、ディスク復旧時間の推定指示248を設定し、推定結果249が設定されたら、それをサービスプロセッサ4に通知する。第二に、サービスプロセッサ4よりコピー制御情報(復旧対象ディスクドライブ、コピー速度)の指定を受けたら、コピー開始指示244とコピー制御情報245を設定する。第三に、コピー実行状況246にコピー処理完了が設定されたら、ディスク復旧処理完了をサービスプロセッサ4に通知する。

【0040】サービスプロセッサ4は、操作員がコンソール装置5を介して入力した指示をディスク制御装置2に伝えるとともに、ディスク制御装置2から送られてきた情報をコンソール装置5に表示する。

【0041】図7は、上述の各機能を用いたディスク復旧処理のシーケンス図である。サービスプロセッサ4は、操作員からディスク復旧指示を受け取ると(ステップ3000)、これをディスク制御装置2に通知する(ステップ3010)。

【0042】これに対して、ディスク制御装置2では、指標推定部236で、各コピー速度に対するディスク復旧時間を推定した後(ステップ3020)、それをサービスプロセッサ4に報告する(ステップ3030)。サービスプロセッサ4では、各コピー速度に対するディスク復旧時間の推定結果をコンソール装置5に表示する(ステップ3040)。そして、操作員が所望のコピー速度を指定すると、それをディスク制御装置2に通知する(ステップ3050)。

【0043】ディスク制御装置2では、通知されたコピー速度に従って、コピー処理部233でコピー処理を行う(ステップ3060)。コピー処理完了後、ディスク復旧処理終了をサービスプロセッサ4に報告する(ステップ3070)。サービスプロセッサ4は、コピー処理

完了のメッセージをコンソール装置5に表示して、操作員に知らせる(ステップ3080)。

【0044】なお、操作員がディスク復旧処理を指示すると同時にコピー速度を指定した場合は、指標推定部236は、指定されたコピー速度に対するディスク復旧時間の推定値を推定結果249に設定する。

【0045】以上の第1実施例によれば、操作員がディスク復旧処理を指示したときに、各コピー速度に対するディスク復旧時間の推定値が表示されるので、例えば2

10 台目のディスクドライブが故障してしまう確率を考慮して操作員がコピー速度を指定できるなど、ディスク復旧作業計画を立てやすくなる効果がある。なお、指標推定部236やモニタ部235の機能をサービスプロセッサ4やホストコンピュータ1に移したり、モニタ部235の機能を有する専用の計測機を設けてもよい。

【0046】-第2実施例-

第1実施例ではディスク復旧時間を推定し表示したが、第2実施例では、複数台(本実施例の場合は2台)のディスクドライブの故障によりデータ復元が不可能になるまでの平均時間を推定し表示する。これは、図2の指標推定部236の機能を次のように変えることによって実現する。指標推定部236では、ディスクアレイシステムの状態遷移図を解析する。本実施例の場合の状態遷移図は、図8のようになる。すなわち、以下の4つの状態と6つの遷移を持つ。

状態S0 : 4台のディスクドライブが正常に稼動中

状態S1 : 1台のディスクドライブが故障中で、残りの3台で正常に稼動中

20 状態S2 : 1台のディスクドライブの復旧処理実行中で、残りの3台でも正常に稼動中

状態S3 : 2台のディスクドライブの故障により、データ復元不可能(システムダウン)

遷移(S0, S1) : 4台のディスクドライブの内の1台が故障

遷移(S1, S2) : 故障したディスクドライブを修理・交換

遷移(S2, S0) : ディスクドライブ復旧処理完了

遷移(S2, S1) : 復旧対象のディスクドライブが故障

40 遷移(S1, S3) : 正常な3台のディスクドライブの内の1台が故障

遷移(S2, S3) : 正常な3台のディスクドライブの内の1台が故障

図8で、各遷移上のラベルは遷移率を表す。入は、1台のディスクドライブの故障率を表す。μは、1台のディスクドライブの修理交換率を表す。Tはディスク復旧時間を表す。故障率入および修理交換率μの値は、予めシステムに与えておく。これらが指数分布に従うことを仮定すると、マルコフ過程の理論により、次の微分方程式

15

が成立する。なお、 $P_i(t)$ は、時刻  $t$  において状態  $S_i$  に存在する確率である。

\* [0047]  
\* [数1]

$$\left. \begin{aligned} \frac{dP_0(t)}{dt} &= -4\lambda P_0(t) + \frac{1}{T} P_2(t) \\ \frac{dP_1(t)}{dt} &= 4\lambda P_0(t) - (3\lambda + \mu) P_1(t) + \lambda P_2(t) \\ \frac{dP_2(t)}{dt} &= \mu P_1(t) - (4\lambda + \frac{1}{T}) P_2(t) \\ \frac{dP_3(t)}{dt} &= 3\lambda P_1(t) + 3\lambda P_2(t) \end{aligned} \right\} \cdots \text{ (数1)}$$

【0048】指標推定部236では、サービスプロセッサ通信部237より推定指示248を受けたとき、モニタ結果247からその時点での入出力要求到着率を読み出し、その入出力要求到着率が維持されたと仮定したときの各コピー速度に対するディスク復旧時間を図5の関係から求める。次に、 $\lambda$ 、 $\mu$ 、 $T$ の値を上記(数1)に代入して微分方程式を解く。そして、データ復元が不可 $\cdots\cdots\cdots$ 20

$$MTTDL = \int_0^\infty P_3(t) dt$$

【0050】各コピー速度と平均時間MTTDLの関係は、第1実施例と同様にコンソール装置5に表示される。

【0051】以上の第2実施例によれば、操作員がディスク復旧処理を指示したときに、各コピー速度に対するデータ復元が不可能になるまでの平均時間MTTDLの推定値が表示されるので、それを考慮して操作員がコピー速度を指定できるなど、ディスク復旧作業計画を立てやすくなる効果がある。

【0052】-第3実施例-

第1実施例ではディスク復旧時間の推定値を表示し、第2実施例ではデータ復元が不可能になるまでの平均時間MTTDLの推定値を表示したが、第3実施例では、要求を受けてから処理を終えるまでの応答時間を推定し表示する。第3実施例の構成は、モニタ部235と指標推定部236の機能以外は、第1実施例および第2実施例と同じである。

【0053】モニタ部235では、チャネル21を通る信号を常時モニタして、ホストコンピュータ1からの入出力要求と、入出力処理部232がホストコンピュータ1に対して行う処理終了報告とをチェックする。そして、入出力要求到着率と、要求を受けてから処理を終えるまでの応答時間を計測する。計測した入出力要求到着率と応答時間は、モニタ結果247に記録される。さらに、モニタ部235では、1回のディスク復旧処理の度に、そのときのコピー速度と一緒に、平均入出力要求到着率と平均応答時間とを、モニタ結果247に記録する。コピー速度と平均入出力要求到着率と平均応答時間

※能になるまでの平均時間MTTDL (Mean Time To Data Lost) を次式によって求め、得られた各コピー速度と平均時間MTTDLの関係を推定結果249に設定する。

【0049】

【数2】

..... (数2)

とは統計処理されて、図9に示すような関係が得られる。

【0054】指標推定部236は、サービスプロセッサ通信部237より平均応答時間の推定指示248を受けたとき、モニタ結果247からその時点での入出力要求到着率を読み出し、その入出力要求到着率が維持されたと仮定したときのコピー速度と平均応答時間との関係を求める。そして、その結果を推定結果249に設定する。ある入出力要求到着率におけるコピー速度と平均応答時間の関係は、図10に示すようになる。

【0055】以上の第3実施例によれば、操作員がディスク復旧処理を指示したときに、各コピー速度に対する応答時間の推定値が表示されるので、ディスク復旧処理中のオンライン処理性能を考慮して操作員がコピー速度を指定でき、ディスク復旧作業計画を立てやすくなる効果がある。

【0056】-第4実施例-

40 図6と図10を比較して分かるように、ディスク復旧時間(信頼性)と平均応答時間(オンライン処理性能)は、コピー速度に対してトレードオフの関係にある。第1実施例～第3実施例では操作員がこれらのバランスを抽象的に考慮してコピー速度を指定していたが、第4実施例では、信頼性とオンライン処理性能のバランスを定量的に表わす平均性能維持時間と呼ぶ指標を推定し表示する。

【0057】ディスクアレイシステムのような縮退・復旧可能なシステムでは、稼動している状態の中に、目標性能を上回って稼動している状態(性能維持状態)と下

回っている状態（性能劣化状態）とがある。そこで、性能維持状態と性能劣化状態とを交互に繰り返して、最終的に性能維持状態に戻れなくなるまでの間に、システムが性能維持状態で機能していた総時間を一つの指標として考える。この指標を平均性能維持時間THP (Total \*  

$$THP = \int_0^{\infty} \sum_{S_i \in SH} P_i(t) dt$$

【0059】モニタ部235では、入出力要求到着率と、前記各システム状態S0, S1, S2, S3における応答時間とディスク復旧時間とをモニタして、モニタ結果247に記録する。

【0060】指標推定部236では、サービスプロセッサ通信部237から平均性能維持時間THPの推定指示248を受けたとき、同時に、目標応答時間を受け取る。この目標応答時間は、操作員が、サービスプロセッサ4に対してディスク復旧を指示したときに、併せて入力する。次に、指標推定部236では、モニタ結果247に基づいて、その時点での入出力要求到着率が維持されたと仮定したときの各システム状態S0, S1, S2, S3における各コピー速度に対する応答時間を推定する。そして、その推定した各コピー速度に対する応答時間と前記目標応答時間とを比較することによって、各システム状態S0, S1, S2, S3を性能維持状態と性能劣化状態のいずれかに分類する。次に、指標推定部236では、モニタ結果247に基づいて、その時点での入出力要求到着率に対するディスク復旧時間を求める。次に、 $\lambda$ ,  $\mu$ ,  $T$ の値を上記(数1)に代入して微分方程式を解く。さらに、その結果と前記分類結果とを用いて上記(数3)から各コピー速度に対する平均性能維持時間THPを計算する。

【0061】最後に、指標推定部236では、コピー速度と平均性能維持時間THPの関係を推定結果249に設定する。

【0062】以上の第4実施例によれば、操作員がディスク復旧処理を指示したときに、各コピー速度に対する平均性能維持時間THPの推定値が表示されるので、これから信頼性とオンライン処理性能のバランスを定量的に把握して操作員がコピー速度を指定でき、ディスク復旧作業計画を立てやすくなる効果がある。

【0063】-第5実施例-

第5実施例は、操作員が、ディスク復旧処理中にコピー速度を変更できるようにした実施例である。

【0064】コピー処理部233では、コピー処理が完了したときにコピー実行状況246にコピー処理完了を設定するのに加え、1回のコピー処理が終わる度に、コピーを終えたトラック数をコピー実行状況246に設定する。モニタ部235では、コピー実行状況246を常時チェックして、コピーを終了したトラック数を、コピー開始からの経過時間と一緒に、モニタ結果247に設定する。

\* High Performance Time)と呼ぶ。平均性能維持時間THPは、次式により算出される。なお、SHは、性能維持状態の集合を表す。

【0058】

【数3】

.....(数3)

【0065】指標推定部236では、モニタ結果247から、コピー開始からの実行状況を判断して、コピー処理が完了するまでの残り時間を推定する。そして、推定した残り時間を推定結果249に設定する。この処理は、常時、自発的に行ってもよいし、操作員からの要求があったことを推定指示248を参照することによって検出した時に行ってもよい。

【0066】サービスプロセッサ通信部237では、推定結果249をサービスプロセッサ4に通知するとともに、操作員から改めてコピー速度を指示されたら、それをコピー制御情報245に設定する。コピー処理部233では、コピー処理実行権243を得て、コピー処理を実行する度に、コピー制御情報245からコピー速度を読み込んで、その指示されたコピー速度でコピー処理を実行する。

【0067】以上の第5実施例によれば、ディスク復旧処理中に、コピー処理が完了するまでの残り時間が表示されると共にコピー速度を変更できるので、ディスク復旧処理中に入出力要求到着率が変化するなどした場合でも、操作員が臨機応変に対処することができる効果がある。

【0068】-第6実施例-

第6実施例は、応答時間をモニタして、コピー速度を自動調整する実施例である。図11は、第6実施例のディスクアレイシステム100Bを示す構成図である。モニタ部235では、応答時間を常時計測して、モニタ結果247に設定する。サービスプロセッサ通信部237では、サービスプロセッサ4からディスク復旧指示を受けたとき、併せて目標応答時間を受け取る。そして、その目標応答時間を目標値250に設定するとともに、コピー開始指示244を設定する。

【0069】コピー調整部238では、モニタ結果247から応答時間を常時読み出して、それを目標値250と比較する。目標値の方が大きい場合は、もっとコピー処理を優先することが可能なので、コピー制御情報245に、より大きなコピー速度を設定する。逆に、目標値の方が小さい場合は、コピー処理を抑える必要があるので、コピー制御情報245に、より小さなコピー速度を設定する。

【0070】以上の第6実施例では、ディスク復旧処理中に入出力要求到着率が変化しても、操作員の介入なしに、コピー処理を自動調整して、目標性能を維持することができる。なお、コピー速度を変える外に、ディスパ

ツチ制御部234でコピー処理実行権243を設定する間隔を変えても、コピー処理の優先度を変えることが出来るので、上記と同じ効果が得られる。従って、コピー処理実行権243を設定する時間間隔を、コピー調整部238が自動変更するようにしてもよい。

【0071】-第7実施例-

第7実施例は、平均性能維持時間THPを利用して、コピー速度を自動設定／自動調整する実施例である。図12は、第7実施例のディスクアレイシステム100Cを示す構成図である。まず、第4実施例と同様にして各コピー速度に対する平均性能維持時間THPを計算し、平均性能維持時間THPが最大になるコピー速度を自動設定して、コピー処理を開始する。

【0072】モニタ部235では、常時、応答時間を計測と共に、コピーを終了したトラック数とコピー開始からの経過時間を計測して、モニタ結果247に設定する。指標推定部236では、モニタ結果247から読み出した応答時間を目標応答時間（目標値250）と比較することによって、ディスク復旧処理実行中の現在の状態を性能維持状態か性能劣化状態かに分類する。また、モニタ結果247に基づいて、ディスク復旧時間Tを再計算する。次に、その再計算したディスク復旧時間Tを用いて、平均性能維持時間THPを再計算する。そして、その平均性能維持時間THPを推定結果249に設定する。

【0073】コピー調整部238では、適当な周期で推定結果249を参照し、前回よりも平均性能維持時間THPの数値が悪化している場合には、コピー速度を変えたときの平均性能維持時間THPを再推定するように、指標推定部236に指示を出す。指標推定部236では、改めて各コピー速度に対する平均性能維持時間THPを再推定して、推定結果249に設定する。コピー調整部238では、この推定結果249をもとに、平均性能維持時間THPが最大になるコピー速度を求めて、コピー制御情報245に設定する。

【0074】以上の第7実施例では、ディスク復旧処理中に入出力要求到着率が変化しても、操作員の介入なしに、信頼性とオンライン処理性能のバランスが最適になるように、コピー速度を自動調整できる。

【0075】なお、システムの能力に応じて、第7実施例を次のいずれかに変更してもよい。

①コピー調整部238は、適当な周期で推定結果249を参照し、コピー処理開始時に計算した平均性能維持時間THPよりも平均性能維持時間THPの数値が悪化している場合に、コピー速度を変えたときの平均性能維持時間THPを再推定するように、指標推定部236に指示を出す。

②指標推定部236は、モニタ結果247から読み出した応答時間を目標応答時間（目標値250）と比較することによって、ディスク復旧処理実行中の現在の状態を

性能維持状態か性能劣化状態かに分類し、それを推定結果249に設定する。コピー調整部238は、適当な周期で推定結果249を参照し、性能劣化状態になった場合に、コピー速度を変えたときの平均性能維持時間THPを再推定するように、指標推定部236に指示を出す。

③指標推定部236は、適当な周期でモニタ結果247を参照し、ディスク復旧時間Tを再計算する。次に、その再計算したディスク復旧時間Tを用いて、各コピー速度に対する平均性能維持時間THPを再計算する。そして、その平均性能維持時間THPを推定結果249に設定する。

【0076】-第8実施例-

サービスプロセッサをホストコンピュータ1に接続して、操作員からのディスク復旧開始指示が、ホストコンピュータ1からサービスプロセッサを経由して、ディスク制御装置2に届く構成とすれば、操作員がホストコンピュータ1からディスク復旧開始を指示することが出来る。また、ホストコンピュータ1にサービスプロセッサの機能を設けた構成としても、操作員がホストコンピュータ1からディスク復旧開始を指示することが出来る。

【0077】-第9実施例-

第1実施例～第8実施例では、故障したディスクドライブを修理・交換してから、操作員がディスク復旧開始を指示しているが、第9実施例では、ディスクドライブの修理・交換を待たずに自動的にディスク復旧処理を開始する。図13は、第9実施例のディスクアレイシステム100Dを示す構成図である。ディスクアレイシステム100Dは、5台のディスクドライブ31～35を有している。このうちのディスクドライブ35は、スペアのディスクドライブであり、システムが正常に稼動しているときにはアクセスされない。

【0078】また、ディスク制御装置2には、ディスク障害検出部239が設けられている。このディスク障害検出部239は、入出力処理実行中のディスク障害を検出する機能を持ち、ディスク障害を検出すると、コピー開始指示244を設定するとともに、故障したディスクドライブ番号を障害情報251に書き込む。

【0079】コピー調整部238では、コピー開始指示244が設定されたとき、目標値250に予め登録されている目標応答時間とモニタ結果247とを参照して、第6実施例と同様に、コピー制御情報245を設定する。コピー処理部233では、障害情報251とコピー制御情報245とをもとに、ディスクドライブ31～34の内の正常な3台のディスクからデータを読み込み、スペアドライブ35に対して、復元したデータを書き込む。

【0080】以上の第9実施例では、1台のディスクドライブが故障しても、そのディスクドライブの修理・交換を待たずに自動的にディスク復旧処理を開始し、予め

21

設定された目標性能を守りながらディスク復旧処理を実行することが出来る。なお、応答時間の代わりに、第7実施例のように平均性能維持時間THPを用いてディスク復旧処理を実行してもよい。

## 【0081】

【発明の効果】本発明のディスクアレイシステムによれば、ディスク復旧処理を行う際に、コピー処理をどの程度優先すべきかを決定するための指標を計算して、操作員に提示したり、それに基づいて最適な制御を行うことが出来る。従って、信頼性とオンライン処理性能の両方の観点から適切なディスク復旧処理を行うことが出来るようになる。

## 【図面の簡単な説明】

【図1】本発明の実施例にかかるディスクアレイシステムの構成図である。

【図2】本発明の第1実施例のディスクアレイシステムの構成図である。

## 【図3】コピー処理のフローチャートである。

## 【図4】スケジュール処理のフローチャートである。

【図5】平均入出力要求到着率、コピー速度、ディスク復旧時間の関係図である。

【図6】コピー速度とディスク復旧時間の関係図である。

【図7】ディスク復旧処理のシーケンスのフロー図である。

10

20

22

【図8】図1のディスクアレイシステムの状態遷移図である。

【図9】平均入出力要求到着率、コピー速度、平均応答時間の関係図である。

【図10】コピー速度と平均応答時間の関係図である。

【図11】本発明の第6実施例のディスクアレイシステムの構成図である。

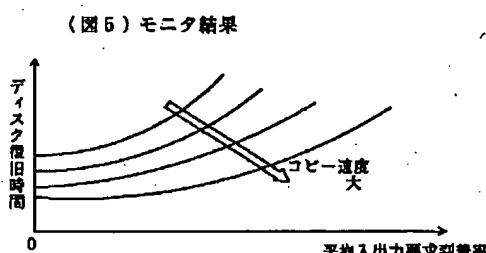
【図12】本発明の第7実施例のディスクアレイシステムの構成図である。

【図13】本発明の第9実施例のディスクアレイシステムの構成図である。

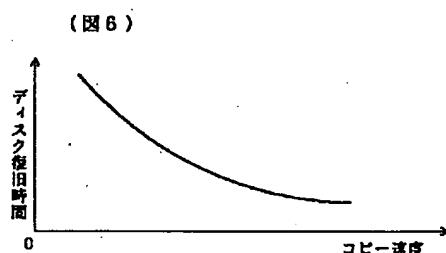
## 【符号の説明】

1…ホストコンピュータ、2…ディスク制御装置、3 1～3 5…ディスクドライブ、4…サービスプロセッサ、5…コンソール装置、2 1…チャネル、2 2…バッファメモリ、2 3 1…入出力要求受諾部、2 3 2…入出力処理部、2 3 3…コピー処理部、2 3 4…ディスパッチ制御部、2 3 5…モニタ部、2 3 6…指標推定期、2 3 7…サービスプロセッサ通信部、2 3 8…コピー調整部、2 3 9…ディスク障害検出部、2 4 1…入出力処理待ちキュー、2 4 2…入出力処理実行権、2 4 3…コピー処理実行権、2 4 4…コピー開始指示、2 4 5…コピー制御情報、2 4 6…コピー実行状況、2 4 7…モニタ結果、2 4 8…推定指示、2 4 9…推定結果、2 5 0…目標値、2 5 1…障害情報。

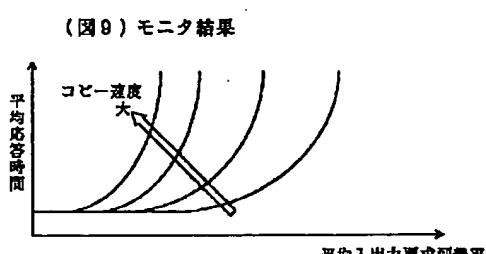
【図5】



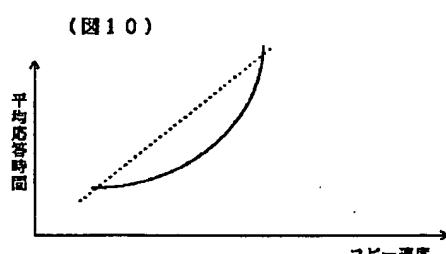
【図6】



【図9】

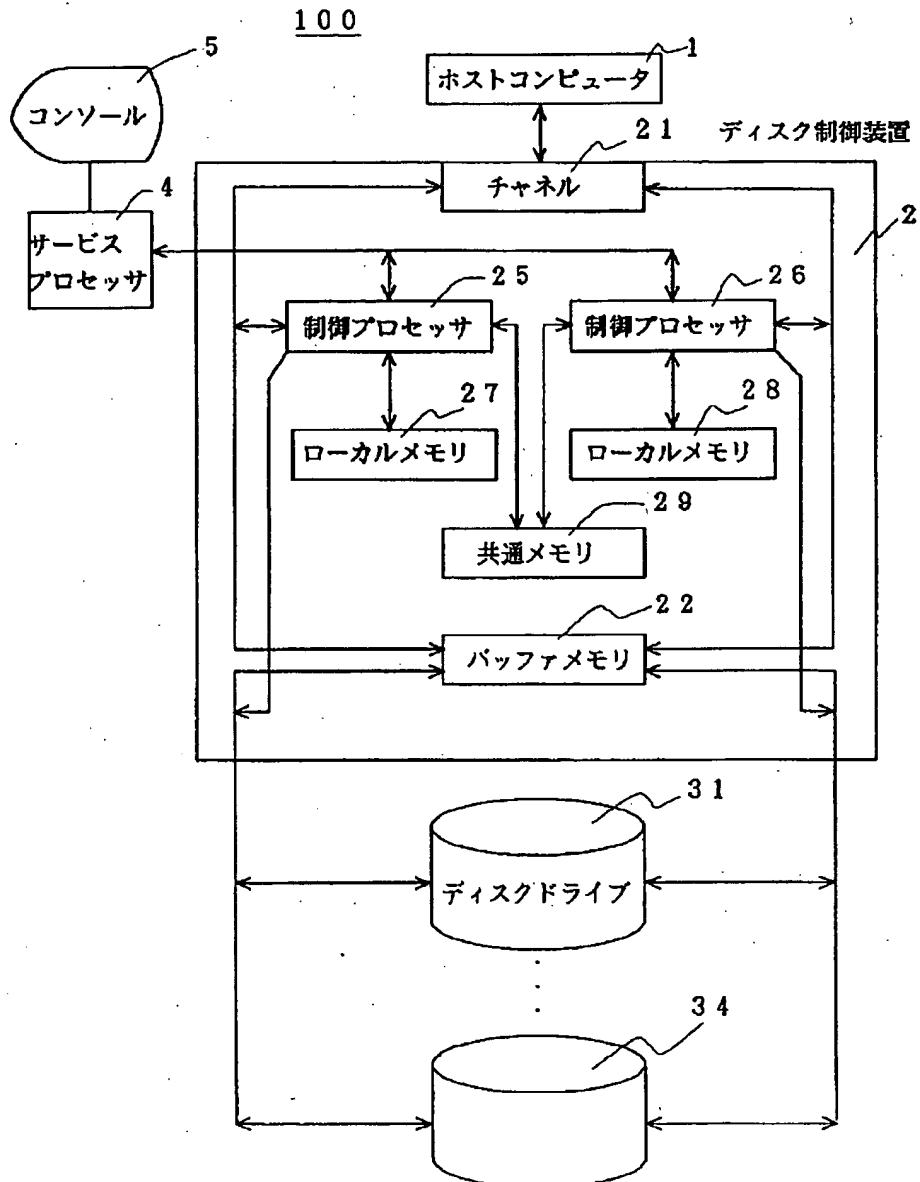


【図10】



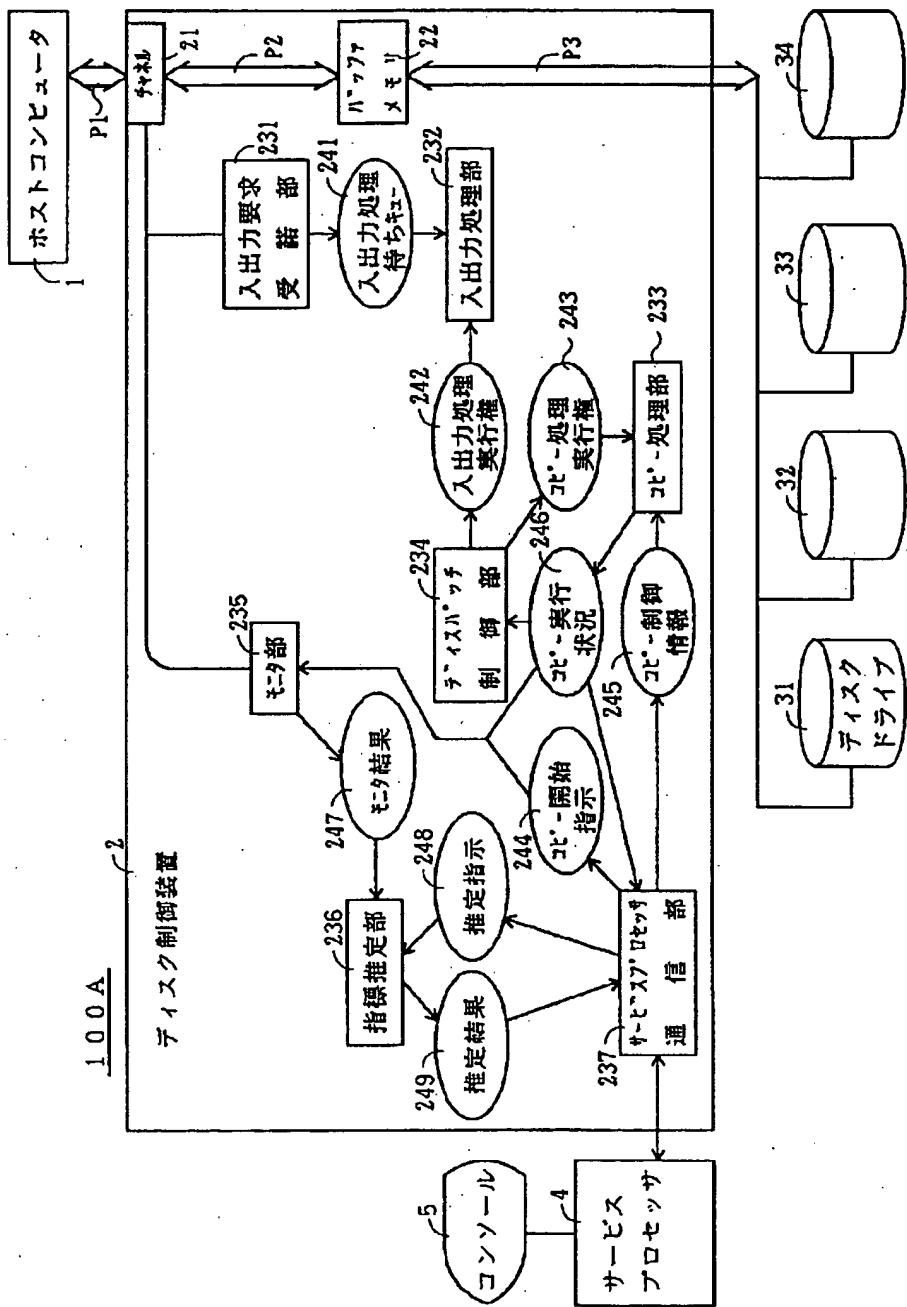
【図1】

(図1)



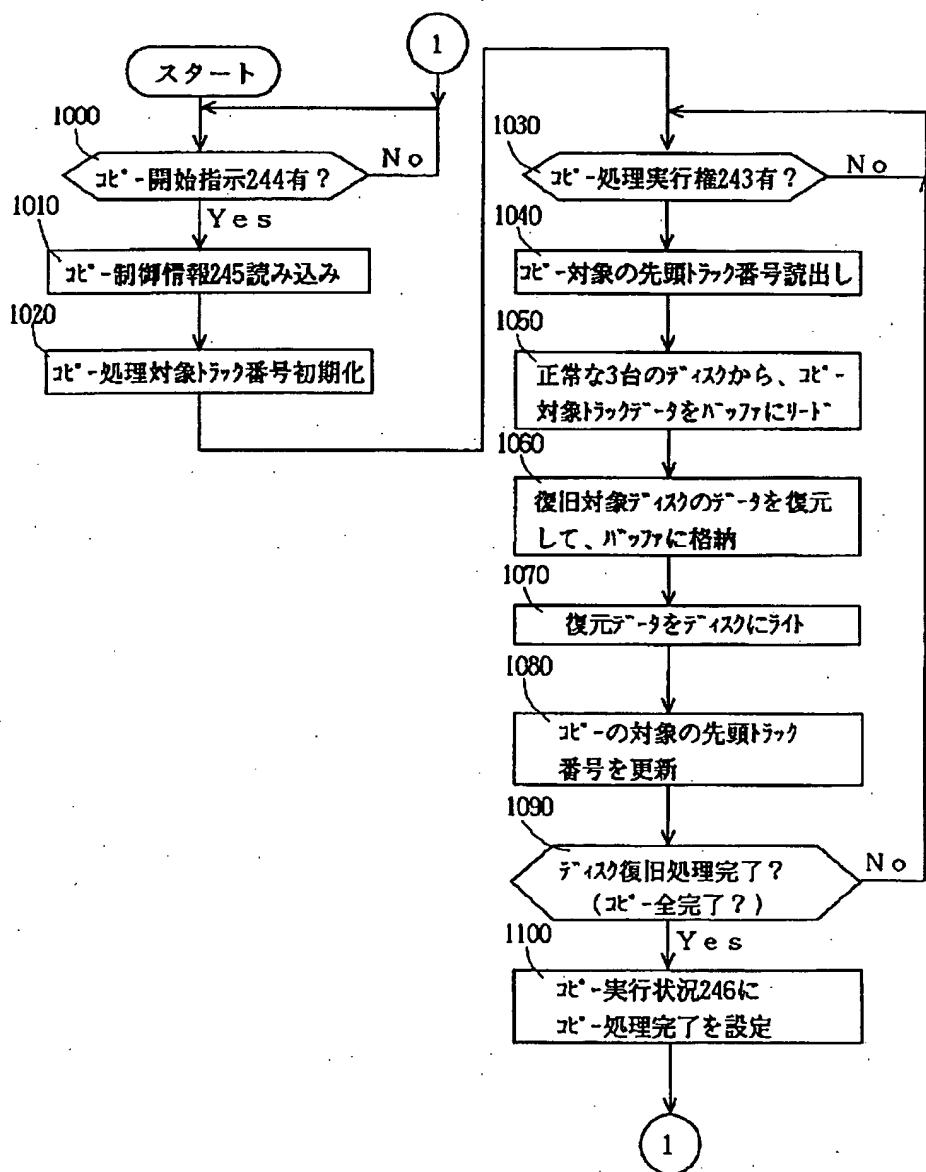
【図2】

{☒ 2}



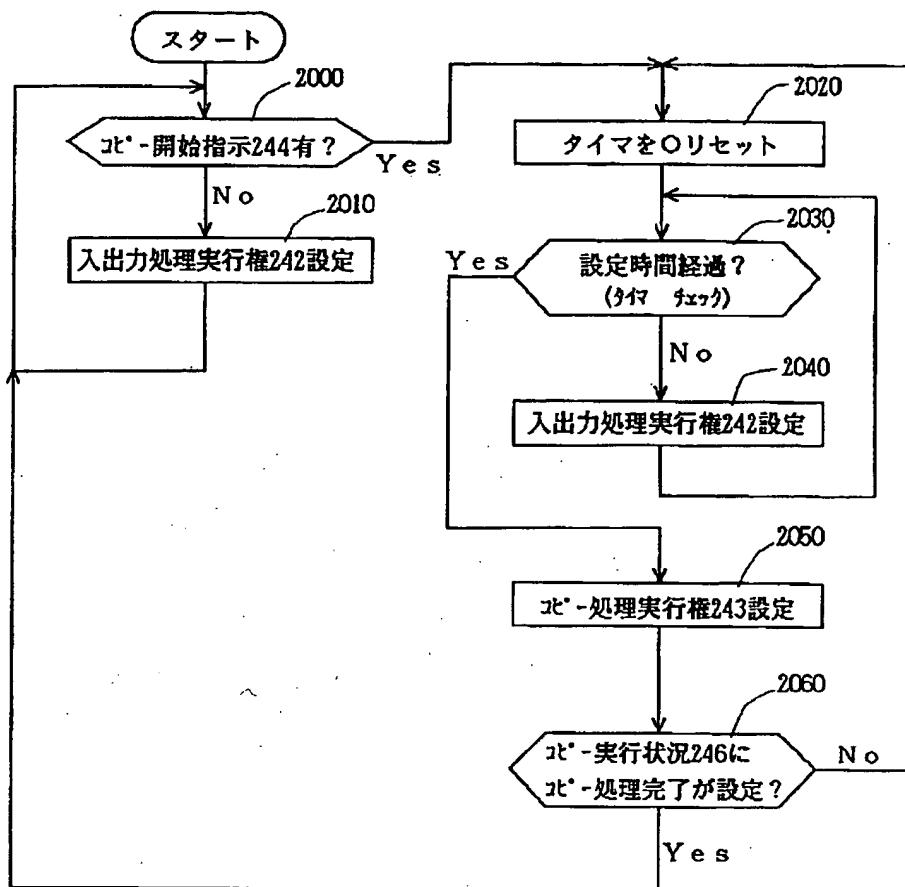
【図3】

コピー処理(図3)



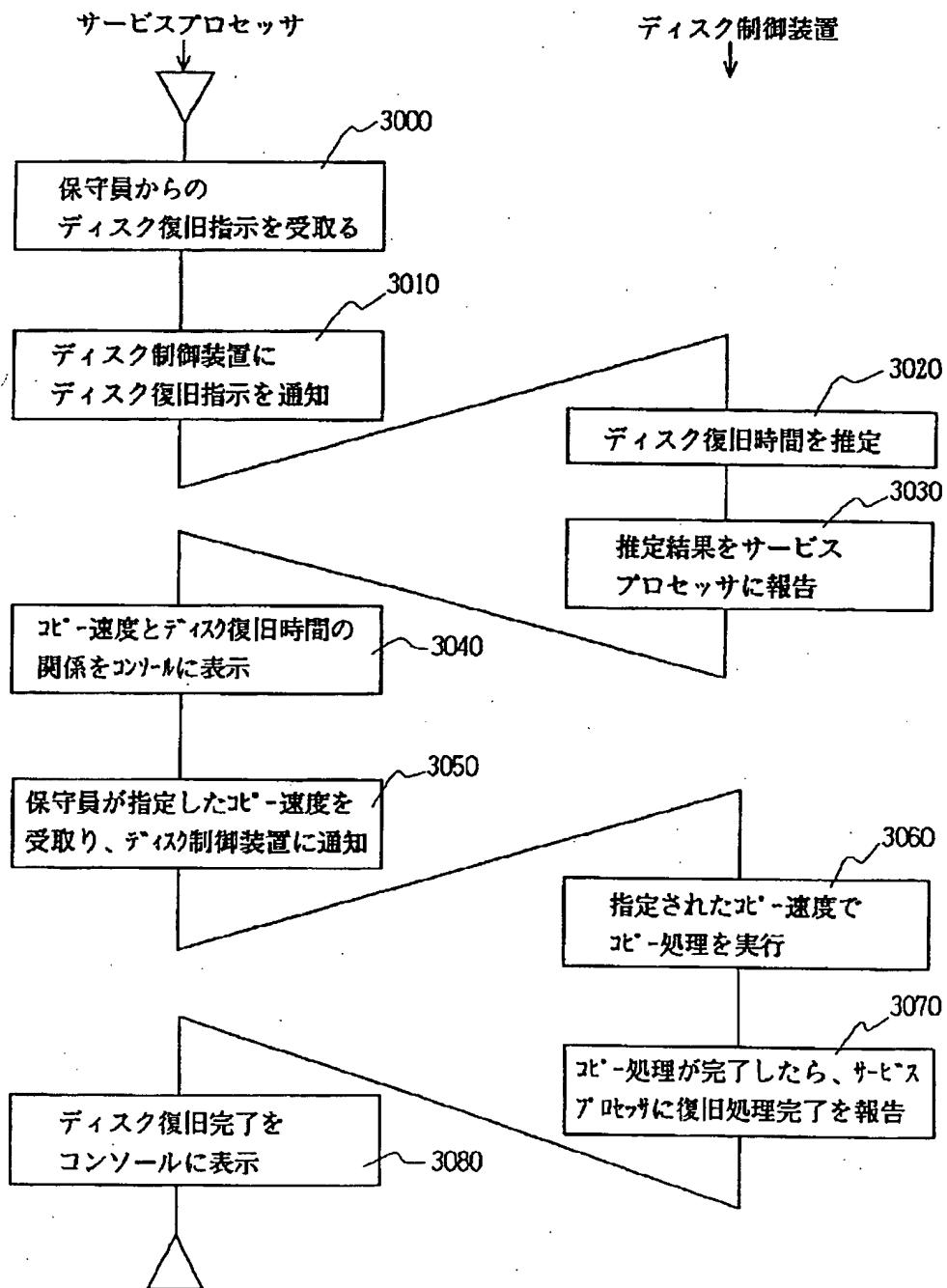
【図4】

スケジュール処理(図4)



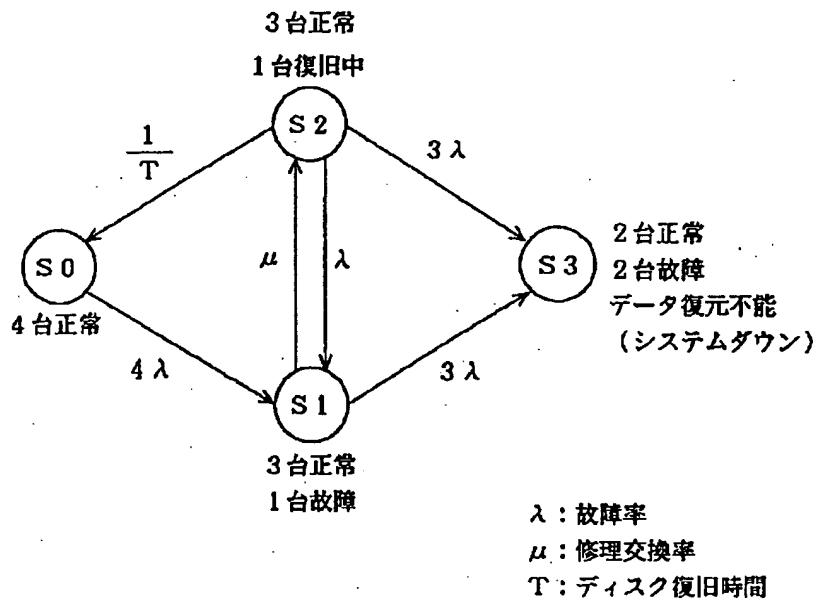
【図7】

## ディスク復旧処理(図7)



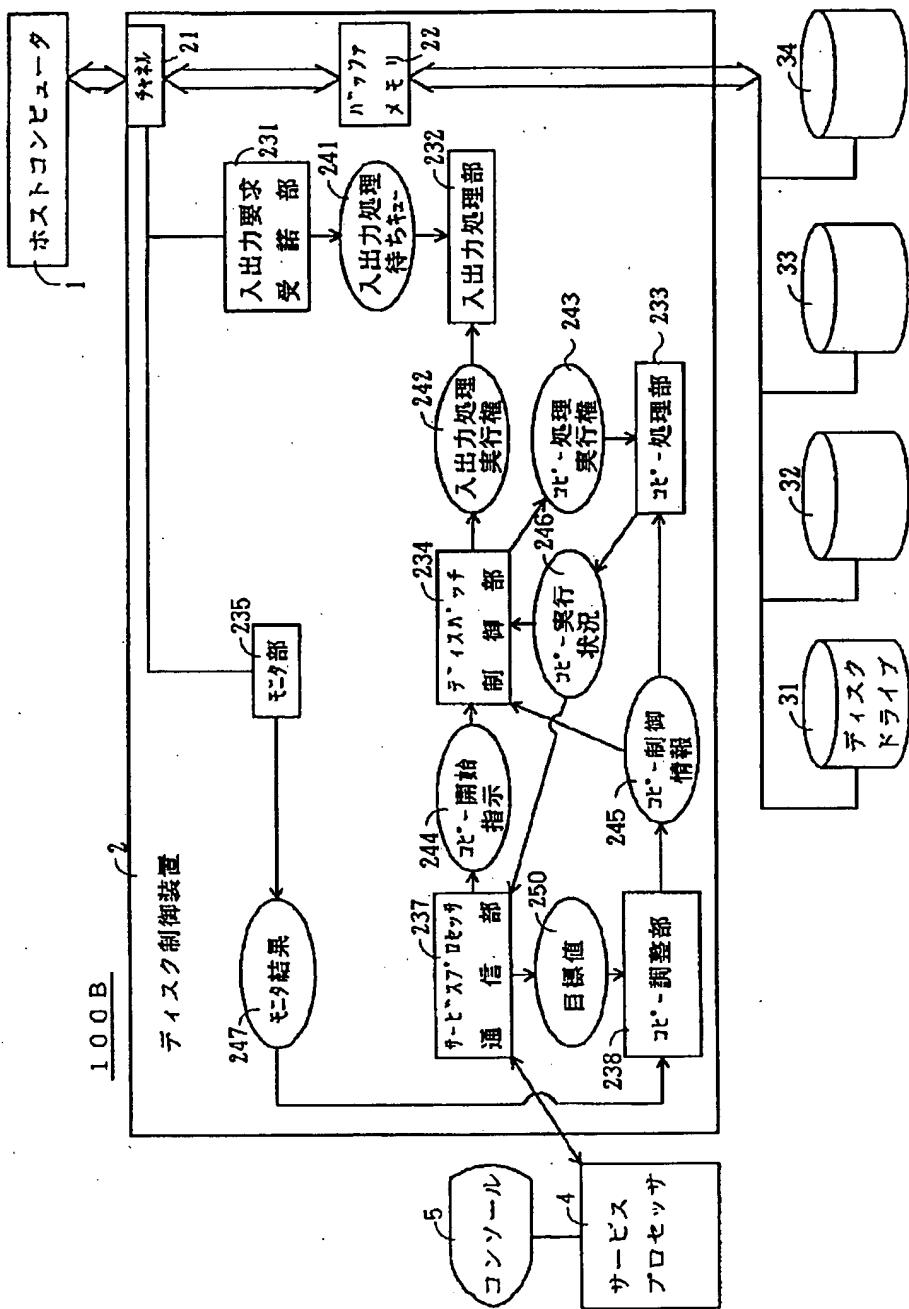
【図8】

システム状態遷移図(図8)



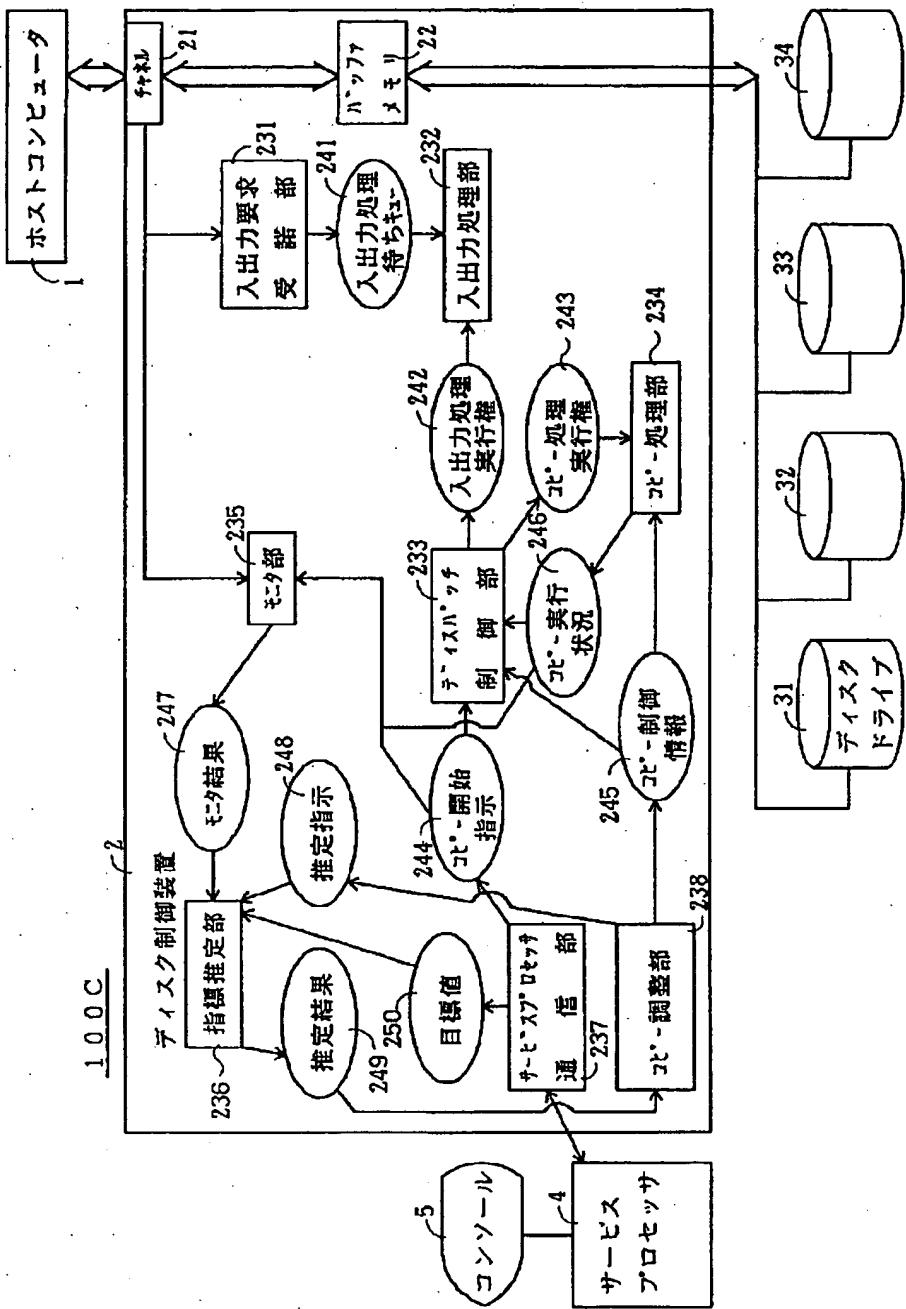
【図11】

【図11】

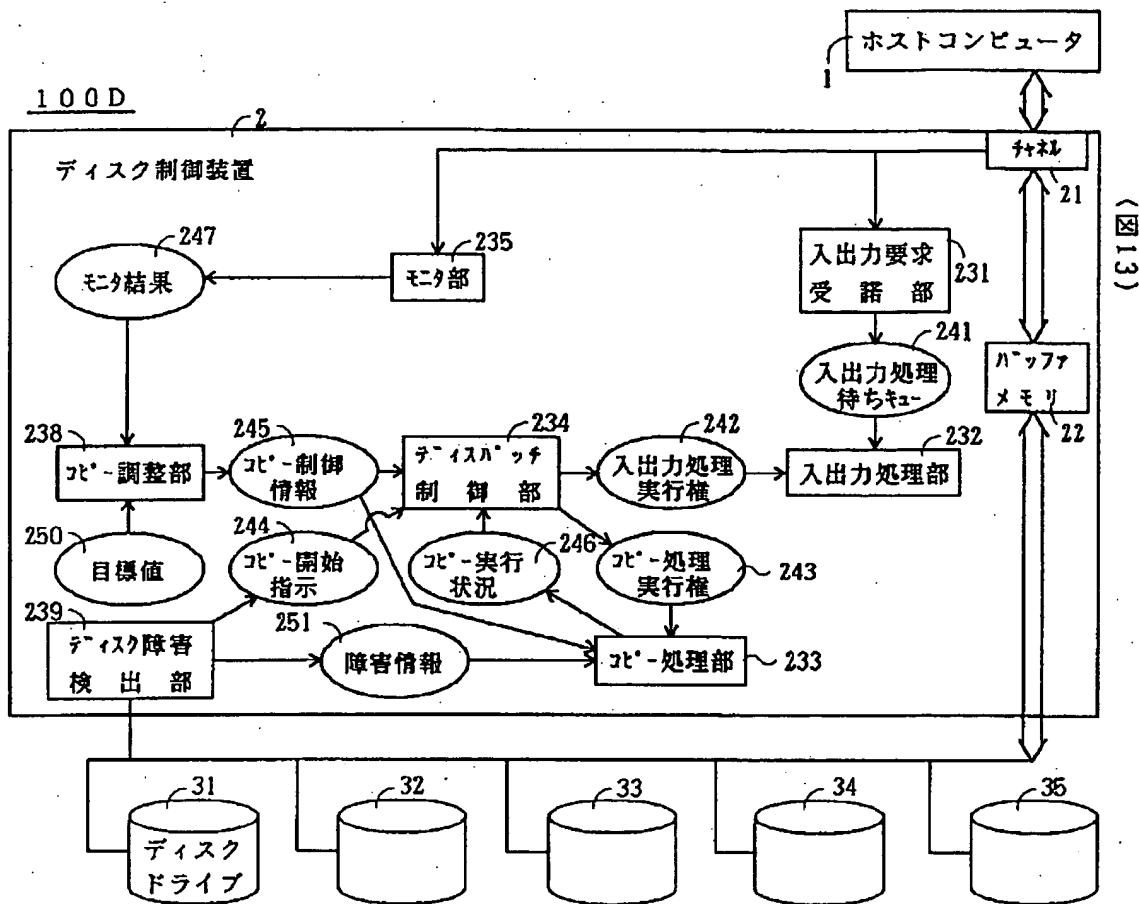


【図12】

(図12)



[図13]



フロントページの続き

(72)発明者 倉野 昭

神奈川県小田原市国府津2880番地 株式会  
社日立製作所ストレージシステム事業部内